AD_____

Award Number: DAMD17-01-1-0463

TITLE: Outcome Based Screening for Prognostic Phospho-RTK (Receptor Tyrosine Kinase) Antibodies Using Tissue Microarrays

PRINCIPAL INVESTIGATOR: David L. Rimm, M.D., Ph.D.

CONTRACTING ORGANIZATION: Yale University
New Haven, Connecticut 06520-8047

REPORT DATE: August 2002

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 2002	Annual (1 Aug 01 - 31 Jul 02)

4. TITLE AND SUBTITLE
Outcome Based Screening for Prognostic Phospho-RTK (Receptor Tyrosine Kinase) Antibodies Using Tissue Microarrays

5. FUNDING NUMBERS
DAMD17-01-1-0463

6. AUTHOR(S)
David L. Rimm, M.D., Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Yale University
New Haven, Connecticut 06520-8047

E-Mail: david.rimm@yale.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
report contains color

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. Abstract *(Maximum 200 Words)* *(abstract should contain no proprietary or confidential information)*

Receptor Tyrosine Kinases (RTKs) have been identified as potential targets for both breast cancer prognosis and therapy. We proposed use of tissue microarrays to evaluate the prognostic value of RTKs with emphasis on the phosphorylation status of these receptors. Analysis of a series of phospho-receptor antibodies on large cohorts on tissue microarrays should reveal which RTKs are most likely to be of prognostic and therapeutic value. As of this progress report, we have completed construction of the tissue microarrays and completed collection of the clinical data. We have also completed and submitted a pilot study of RTKs using conventional analysis of this array. In the proposal we show automated analysis of the arrays has potential to reveal relationships that are undetectable by conventional methods. We have now completed our efforts in construction of the device capable of high through-put automated array analysis. We anticipate testing phospho-RTK antibodies using this device within the next year.

14. SUBJECT TERMS
breast cancer, RTKs, tissue microarrays

15. NUMBER OF PAGES
49

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Introduction:

Biologically specific therapies represent a great new hope for combating cancer. Perhaps the best example of this is Herceptin, a specific drug for a subset of patients with HER2 positive breast cancer. Unfortunately, the pathway from initial discovery to clinical usage is long and slow, taking over 10 years for HER2 (a member of the receptor tyrosine kinases (RTK) family. In this study we proposed a potential method for rapid evaluation of other RTKs as bio-specific therapies based on prediction of outcome. We proposed that phospho-specific antibodies to RTK would provide highly specific prognostic markers for outcome and predictive markers for response to anti-receptor type therapies. Our objective was to produce specific phospho-RTK antibodies and then to evaluate them using breast cancer tissue microarrays. Breast cancer tissue microarrays are a method of placing 0.6 mm diameter samples of breast cancers from hundreds of patients on a single slide. Using this method we proposed high throughput screening for potential prognostic antibodies at the earliest stages of antibody development.

Body:

The main difficulty in finding these potential new RTK targets is that there are many potential candidates and it is difficult and expensive to evaluate a large number of antibodies on large breast cancer cohorts using conventional methods. The novelty or "idea" of this proposal is the use of a newly described, high throughput mechanism for evaluation of antibodies. Thus instead of producing candidate antibodies and screening using the conventional approach (which can take many years), we propose reversal of this process, screening for candidates using the outcome testing, then only proceeding with development of antibodies that have already been validated on large populations of breast cancer patients. We originally proposed the following specific aims:

1. Construction of a series of candidate phospho-RTK antigens and immunization into mice.
2. Evaluation of mouse test bleeds using 250 case, 3-fold redundant, breast cancer cohort tissue microarrays.
3. Selection of promising sera and production of monoclonal antibodies followed by confirmatory testing on tissue microarrays.

Our original Statement of Work was as follows:

Year 1:

1. Select and produce phospho- and corresponding non-phospho-peptides representing critical sequences of RTKs
2. Select 250 breast cancer cases and begin construction of tissue microarray
3. Inoculate mice with first set of phospho-peptides and collect test bleeds.
4. Begin first array screening

Year 2

1. Collect test bleeds and screen arrays.
2. Do final inoculations and test bleeds
3. Begin full scale TMA verification of first promising candidates
4. Do fusions and begin production of first Mabs

5. Production of small scale cloning verification tissue microarrays

Year 3

1. Complete production of Mabs
2. Complete verification of Mabs using first small scale TMAs
3. Production and large cohort testing (750+ case) of new antibodies

This progress report describes work completed in year 1.

In evaluation of the optimal RTKs for antibody preparation we consulted many sources. We found that many biotech antibody companies had already embarked on production of phospho-specific and other RTK antibody production. Rather that try to duplicate there efforts, we decided to begin by purchasing some of the 100's of RTK antibodies that are now commercially available.

Prior to testing any antibodies, arrays needed to be constructed and cohort data collected. That process was completed during the first year. We now have constructed 3 master array blocks for analysis. They include YTMA10 , including 350 node positive cases, YTMA 12, including 350 node negative cases, and YTMA 23, including a total of 250 cases representing a subset of the previous cohorts (125 node negative and 125 node positive) that have ample tissue available for analysis of multiple antibodies.

After completion of construction of the arrays, we began a two-fold approach to analysis. As a proof of concept, we selected 5 RTKs (HER2, MET, EGFR, FGFR, and IGFR) as test cases. We analyzed expression of these using YTMA 12 and have completed and submitted for publication the results of this effort. That submission in included in the appendix. In summary, we found that only Met was predictive of outcome in the node-negative population but that we were able to group MET and FGFR one group that was unrelated to HER2 and EGFR which comprised a second, unrelated group. Although we were pleased with these results, the original grant discussed analysis of arrays using a more quantitative method than the ordinal subjective scale used in this study. Toward that aim, other investigators in the lab have finalized a method for automated quantitative analysis of tissue microarrays. This work will be extremely valuable in optimizing the high throughput analysis of arrays in the future. Although the automated analysis development work was not funded by this grant, it impacts the grant dramatically. We have included a manuscript (in revision) in the appendix that describes the new automated analysis system (AQUA).

The automated analysis system was designed on the Applied Precision Deltavision platform, but that platform is not well suited to the task and the system at our institution is very heavily used allowing only limited time for our applications. We have just completed a rebudgeting request and will use the Army grant money for purchase of a new dedicated 2nd generation automated

tissue microarray analysis device. We anticipate receipt of the new device in September. We anticipate that is will be fully operational by mid October.

The next step will be the evaluation of multiple phospho-specific antibodies. We have begun discussions with Cell Signaling Technologies (CST) since they have produced hundreds of phospho-specific antibodies, including many that are against the same RTK sites proposed in the original grant. Instead of duplicating this effort, we will collaborate with CST or purchase CST antibodies for the initial phases of automated analysis of tissue microarrays, currently planned for the next year.

Key Research Accomplishments:

- Completion of technology for high throughput automated analysis of tissue microarrays
- Completion of array construction of both the 250 case array and the 700 case array and cohort
- Completion of RTK conventional pathologist-based analysis of array with node negative cohort
- Testing of phospho-specific antibodies on arrays

Reportable Outcomes:

Analysis of RTKs shows MET is predictive of outcome in a node-negative population and it identifies a series of patients unique from that identified by over-expression of HER2 or EGFR.

Conclusions:

Tissue microarrays are a valuable tool for analysis of protein and phospho-protein expression. Completion of our automated analysis system will allow high throughput analysis of these arrays. We anticipate evaluation of the first series of phospho-specific RTK antibodies within the next year.

The significance of these developments are that we are now well positioned to find a series of valuable prognostic markers. These markers will are likely to be valuable for prognostication, but they are also likely to be valuable in that they have the potential to identify pathways that will be good targets for bio-specific therapeutics.

References:

See original proposal.

Appendices:

1. Camp et al, *Automated Subcellular Localization And Quantification Of Protein Expression In Tissue Microarrays.* Nature Medicine, Submitted
2. Ocal et al, *Tissue Microarray based studies of node-negative breast cancer patients show Met expression associated with worse outcome but not correlated with EGF family receptors.* Cancer. Submitted
3. Camp et al. *Quantitative analysis of breast cancer tissue microarrays shows both high and normal levels of HER2 expression are associated with poor outcome.* In preparation

Automated Subcellular Localization And Quantification Of Protein Expression In Tissue Microarrays

Robert L. Camp[1], Gina Chung[1], and David L. Rimm[1*]

[1]Department of Pathology Yale University School of Medicine, New Haven, CT. 310 Cedar St., New Haven, CT 06510

*To whom correspondence should be addressed, email: david.rimm@yale.edu

The recent development of tissue microarrays – composed of hundreds of tissue sections from different tumors arrayed on a single glass slide – facilitates rapid evaluation of large-scale outcome studies. Realization of this potential depends on the ability to rapidly and precisely quantify the protein expression within each tissue spot. We have developed a set of algorithms that allow the rapid, automated, continuous and quantitative analysis of tissue microarrays, including the separation of tumor from stromal elements and the sub-cellular localization of signals. Validation studies using estrogen receptor in breast carcinoma show that automated analysis matches or exceeds the results of conventional pathologist-based scoring. Automated analysis and sub-cellular localization of β-catenin in colon cancer identifies two novel, prognostically significant tumor subsets, not detected by traditional pathologist-based scoring. Development of automated analysis technology empowers tissue microarrays for use in discovery-type experiments (more typical of cDNA micro-arrays) with the added advantage of inclusion of long-term demographic and patient outcome information.

Despite the promise of automated analysis of histologic sections, it has failed to replace traditional, pathologist-based evaluation, even in the simplest of conditions such as the analysis of immunohistochemical stains. Whereas the automated analysis of isolated cells in fluids or smears (e.g. fluorescent cell sorting and laser scan cytometry) is now routine [1], the analysis of tissue sections is hampered by the fact that tumor tissue is a complex mixture of overlapping malignant tumor cells, benign host-derived cells and extracellular material. Several methods (including confocal and convolution/deconvolution microscopy) can determine the sub-cellular localization of target antigens, but only through computationally intensive techniques, requiring the acquisition of multiple high power serial images. [2]. Methods designed for tissue microarrays perform only limited subcellular localization using morphometry and usually require significant manual interface (e.g. drawing polygons around tumor cells)[3,4]. In general, pathologist-based analysis remains the current standard for the immunohistochemical studies.

Tissue microarrays provide a high-throughput method of analyzing the prognostic benefit of a myriad of potential targets on large cohorts of patient samples [5-7], but are limited by the pathologist's ability to reproducibly score on a continuous

scale, discriminate between subtle low-level staining differences, and accurately score expression within sub-cellular compartments. We have developed a set of algorithms that we call AQUA (Automated Quantitative Analysis) that allow the rapid, automated analysis of large-scale cohorts on tissue microarrays. The first, called PLACE (Pixel-based Locale Assignment for Compartmentalization of Expression) utilizes fluorescent tags to separate tumors from stroma and to define sub-cellular compartments. The distribution of a target antigen is then quantitatively assessed according to its co-localization with these tags. Since sub-cellular compartments (e.g. membrane, cytoplasm, nuclei, etc.) of different tissues and tumors vary widely in size and shape, traditional methods of defining compartments based on morphometric criteria (i.e. feature extraction) perform poorly on a large scale basis. Rather than counting target-containing features, PLACE delineates target expression as the sum of its intensity divided by the total size of the assayed compartment.

Since the thickness of tissue sections makes it difficult to discriminate between overlapping sub-cellular compartments, we have also developed a novel, rapid exponential subtraction algorithm (RESA), which subtracts an out of focus image, collected slightly below the bottom of the tissue, from an in-focus image, based on pixel intensity, signal-to-noise ratio, and the expected compartment size. This algorithm dramatically improves the assignment of pixels to a particular subcellular compartment (Fig. 1). For a more complete discussion of the image manipulations performed in this protocol, see the Supplementary Information or http://www.yalepath.org/dept/research/YCCTMA/tisarray.htm.

Our initial validation of this technology compared its accuracy, intra-observer variability, and predictive power to traditional pathologist-based analysis. We stained a tissue microarray derived from 340 node-positive breast-carcinoma patients for the presence of estrogen receptor -- the oldest and most common prognostic marker for breast cancer [8]. First we analyzed the ability of automated analysis to match results from a pathologist-based evaluation and found a high degree of correlation (R = 0.884, fig 2A). Next, we compared the variability of a pathologist-based and automated analysis of two separate histospots derived from the same tumor (Fig 2B&C). This comparison shows that automated analysis has slightly better reproducibility (R = 0.824 vs. R=0.732).

Although automated analysis compares favorably with pathologist-based interpretation of microarrays, the true criterion standard is outcome prediction. Estrogen receptor expression is known to significantly improve outcome, because it is associated with less aggressive tumors that are more responsive to anti-estrogenics (e.g. Tamoxifen). We compared the survival of patients with tumors with high (top 25%) versus low (bottom 25%) estrogen receptor expression as assessed by both automated and pathologist-based scoring (fig 2D). Results show that both methods provide similar prognostic information (RR = 2.44 vs. 2.06, automated vs. pathologist); although the automated analysis shows slightly higher significance ($P = 0.0003$ vs. $P = 0.0020$). Univariate analysis of the automated analysis shows a relative risk of 2.438 (p=0.0005, 95% CI 1.480-4.016). When analyzed in a multivariate analysis against histologic and nuclear grade, age, and stage, automated ER analysis retains independent prognostic

significance (RR = 2.566, 95% CI 1.428-4.611, p=0.0016). The pathologist-based analysis shows similar results, validating the cohort (see supplemental information). To determine the reproducibility of our automated analysis of ER, we used the "split sample technique," by dividing the cohort into halves and using one half as a "training" set and the other as a "test" set [9]. The training set was used to determine standard cutoffs for the top and bottom 25% of cases. These cutoffs were then used to divide the test set into top, middle, and bottom groups. We analyzed 300 randomly selected training and test sets; on average 97% of the test cases were correctly classified.
One clear advantage to automated analysis, is that it can perform a true continuous assessment of a target. In contrast, the human eye, even that of a trained pathologist, has a difficult time accurately distinguishing subtle differences in staining intensity using a continuous scale. Consequently, scoring systems for pathologists tend to be nominal (e.g. 0, 1+, 2+, 3+). Algorithms such as the "H-score" are meant to translate such nominal observations into semi-quantitative results. However, the inability to detect subtle differences in staining intensity, particularly at the low and high ends of the scale, as well as the tendency to round scores limits the effectiveness of the H-score. The discontinuity of pathologist-based scoring, despite the use of an H-score algorithm, is exemplified in the estrogen-receptor staining results in figure 2. Note the preponderance of scores at 0, 100, 200, and 300. Furthermore, on average, over half of the cases were assigned to one extreme or the other (39% at 0, and 12% at 300). Thus 51% of the cases could not be effectively ranked. In contrast, the range of scores from the automated analysis is continuous from 0 to 1000. We hypothesize that the two key advantages of automated assessment, continuity of scoring and accurate sub-cellular localization, will allow tumor classification beyond that attainable by current methods.

To demonstrate this potential, we analyzed β-catenin expression in colon cancer. β-catenin is an ideal candidate in that it exhibits complex sub-cellular localization and manifests oncogenic properties upon localization to the nucleus [10]. Numerous studies have shown that β-catenin plays a dual role in both cell-cell adhesion and cell proliferation, depending on its location (reviewed [11]). Membrane-associated β-catenin stabilizes cadherins-mediated adhesion by facilitating the cytoskeletal attachment of adhesion complexes. In contrast, nuclear-associated β-catenin activates several genes important in cell proliferation and invasion [12]. In development, translocation of β-catenin to the nucleus results from *wnt*-mediated cell signaling [13]. However, spurious activation of this pathway is often seen in tumors through mutation of β-catenin or other proteins involved in its activation and/or degradation [14]. Studies on the prognostic value of β-catenin have been mixed [15-17].

The complex biology and uncertain prognostic value of β-catenin made it a suitable candidate for assessing the value of quantitative subcellular localization. We studied a cohort of 310 colon cancers, using both pathologist-based and automated systems for scoring overall, nuclear, and membrane-associated levels of β-catenin expression. Manual analysis used a traditional 4-point nominal scale (0 through 3+), while automated analysis used a continuous 1000-point scale. In a previous study using a similar cohort, we were unable to find prognostic value in assessing nuclear β-catenin levels [18]. This data was confirmed in our present study when comparing tumors

expressing the highest levels of nuclear β-catenin (3+, representing 19% of the cases) versus the rest (fig 3A, $P = 0.2354$). We hypothesized that with the benefit of automated, continuous assessment, these 3+ cases could be subdivided into cases expressing very-high versus high levels. We began by analyzing the top 25% of tumors expressing nuclear β-catenin, as assessed by automated analysis. This group shows a trend toward poorer survival (fig 3B, $P = 0.0760$). When we subset the tumors to assess the top 10% expressers, there is a statistically significant survival difference (fig 3C, $P = 0.0332$, relative risk = 1.740). Further fractionation of the data reveals that the top 6.7% (15^{th} percentile) exhibit even poorer survival with a higher statistical significance (fig 3D, $P = 0.0038$, relative risk = 2.415). This analysis demonstrates the power of continuous automated assessment to define subsets of tumors not seen using standard pathologist-based assessment. Unlike beta-catenin, subdividing ER into smaller and smaller subsets does not significantly alter its prognostic ability, suggesting that ER may be a truly continuous marker where no subpopulations exist.

We then attempted a tumor classification based on comparative sub-cellular localization. Since the translocation of β-catenin from the membrane to the nucleus is thought to correlate with transcriptional activation, we analyzed the ratio of nuclear to membrane-localized β-catenin. By its nature, this type of analysis is essentially impossible without continuous scoring. A crude measurement of overall β-catenin levels using either a pathologist-based or an automated system fails to demonstrate a significant difference in survival between the highest and lowest expressing tumors (fig 4A & B, $P = 0.9425$ and $P = 0.4551$, respectively). In contrast, when we ratio the level of nuclear/membrane β-catenin, we find that tumors with a high ratio have a worse outcome than tumors with a low ratio (RR = 1.718, P = 0.0284). Note that this method defines a relatively large subset (25%) of tumors with poor prognosis, the majority of which are not identified by analyzing individual subcellular compartments. Indeed, comparison the tumors with the highest nuclear/membrane ratio (top 25%) versus the highest overall nuclear levels of β-catenin (top 25%), shows that there is only 47% overlap between the two subsets. Multivariate analysis of nuclear/membrane β-catenin ratios shows independent prognostic significance when analyzed with tumor size, nodal status, tumor grade, and patient age (RR = 1.865, 95% CI = 1.068-3.259, P = 0.0285). In contrast, multivariate analysis of total nuclear β-catenin levels fails to show independent prognostic significance because it is highly correlated with nodal metastases (see supplemental information).

The methods presented here are highly adaptable to a number of tumor types and target markers. In most cases, compartment-specific tags are identical regardless of tumor type (DAPI for nuclei, cadherins/catenin complexes for membranes). Because the methods do not use heuristic models that require recognition of compartments according to size, shape, or texture, they are fully adaptable to tumors with overlapping or pleomorphic cells and/or nuclei. Furthermore, the algorithms can be easily expanded to cover novel compartments or even "virtual" compartments (e.g. mitochondria, lysosomes, cortical actin ring) or tumor types (e.g. mesothelioma), as long as tags can be identified for the prospective compartments / cell types. In addition to estrogen receptor and β-catenin, we have used these techniques to successfully analyze dozens of

markers including growth factor receptors, intracellular signaling molecules, and proliferation markers, using both conventional and phospho-specific antibodies (see supplemental information). Analysis of these targets requires only a standard antibody titration.

We have found that most antigens benefit from sub-cellular localization. Localization can be simple, such as determining the amount of the proliferation marker KI-67 in tumor nuclei, or more complex as in the case of β-catenin. Localization of intracellular signaling molecules (e.g. STATs) may be vital in assessing their potential as prognostic markers. In addition, recent studies have shown that membrane-bound growth factor receptors (e.g. Epidermal Growth Factors, EGFR and ERB-B4, and Fibroblast Growth Factor) can translocate to the nucleus and may act as transcriptional regulators [19]. Sub-cellular localization of such markers may be critical to their use as prognostic markers in cancer.

Depending upon the array size, and the complexity of the compartmentalization, analyses using our current device take from 1-3 hours for image acquisition, and 1-2 hours for analysis. In our laboratory, the average pathologist-based analysis rate is 50-100 spots per hour and usually is performed in several sessions. To increase precision, two or more pathologists read the same array independently and then together to resolve discrepancies. Aside from being more accurate and more robust, automated analysis can be performed continuously and results tabulated immediately. We estimate that a fully integrated tissue microarray reader could be 30 to 50 times faster than pathologist-based scoring.

Our data show that quantitative, continuous-scale, compartmentalized automated analysis of tissue microarrays can provide a rapid assessment of prognosis-based subsets in a variety of tumor markers that cannot be attained using pathologist-based techniques. Automated analysis is better able to discern subtle differences in staining intensity, particularly at the upper and lower extremes, which can distinguish novel prognostic associations. Furthermore, analysis of the sub-cellular distribution of certain signals, using the PLACE and RESA algorithms may elucidate previously unrecognized associations with patient survival. The automated nature of this technology can allow high-throughput screening of tissue microarrays, facilitating their use in large scale, high throughput applications such as target discovery and prognostic marker validation. If, someday, diagnostic criteria are based on molecular expression patterns, the digital nature of this analysis could allow a device of this type to make specific molecular diagnoses.

Methods

Tissue microarray design and processing

Paraffin-embedded formalin-fixed specimens from 345 cases of node-positive breast carcinoma (1962-1977) and 310 cases of colon carcinoma (1971-1982) were obtained, as available, from the archives of the Yale University, Department of Pathology. Microarray slides were prepared, processed and stained as described in the Supplemental Information. For manual analysis, slides were visualized with diaminobenzidine (DAB). For automated analysis, slides were visualized with Cy-5 tyramide.

Image and data analysis

Monochromatic images of tissue microarray histospots were obtained using fluorescently-labeled compartment specific tags (anti-cytokeratin, DAPI, alpha-catenin) as well as target signals (ER and beta-catenin). Images were analyzed using RESA and PLACE algorithms as detailed in the Supplementary Information. Overall survival analysis was assessed using Kaplan-Meier analysis and the Mantel-Cox logrank score for assessing statistical significance. Relative risk was assessed using the univariate and multivariate Cox-proportional hazards model. Analyses were performed using Statview 5.0.1 (SAS Institute, Cary NC).

References

1. Tarnok, A. & Gerstner, A.O. Clinical applications of laser scanning cytometry. *Cytometry* **50**, 133-43 (2002).
2. Robinson, J.P. Principles of confocal microscopy. *Methods Cell Biol* **63**, 89-106 (2001).
3. Rao, J., Seligson, D. & Hemstreet, G.P. Protein expression analysis using quantitative fluorescence image analysis on tissue microarray slides. *Biotechniques* **32**, 924-6, 928-30, 932 (2002).
4. Bacus, S. et al. Potential use of image analysis for the evaluation of cellular predicting factors for therapeutic response in breast cancers. *Anal Quant Cytol Histol* **19**, 316-28 (1997).
5. Kallioniemi, O.P., Wagner, U., Kononen, J. & Sauter, G. Tissue microarray technology for high-throughput molecular profiling of cancer. *Hum Mol Genet* **10**, 657-62 (2001).
6. Kononen, J. et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens [see comments]. *Nat Med* **4**, 844-7 (1998).
7. Rimm, D.L. et al. Tissue microarray: a new technology for amplification of tissue resources. *Cancer J* **7**, 24-31 (2001).
8. Osborne, C.K. et al. Estrogen receptor, a marker for human breast cancer differentiation and patient prognosis. *Adv Exp Med Biol* **138**, 377-85 (1981).
9. Wasson, J.H., Sox, H.C., Neff, R.K. & Goldman, L. Clinical prediction rules. Applications and methodological standards. *N Engl J Med* **313**, 793-9 (1985).
10. Kobayashi, M. et al. Nuclear translocation of beta-catenin in colorectal cancer. *Br J Cancer* **82**, 1689-93 (2000).
11. Provost, E. & Rimm, D.L. Controversies at the cytoplasmic face of the cadherin-based adhesion complex. *Curr Opin Cell Biol* **11**, 567-72 (1999).
12. Morin, P.J. beta-catenin signaling and cancer. *Bioessays* **21**, 1021-30 (1999).
13. Peifer, M. & Polakis, P. Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus. *Science* **287**, 1606-9 (2000).
14. Wong, C.M., Fan, S.T. & Ng, I.O. beta-Catenin mutation and overexpression in hepatocellular carcinoma: clinicopathologic and prognostic significance. *Cancer* **92**, 136-45 (2001).
15. Gunther, K. et al. Predictive value of nuclear beta-catenin expression for the occurrence of distant metastases in rectal cancer. *Dis Colon Rectum* **41**, 1256-61 (1998).
16. Maruyama, K. et al. Cytoplasmic beta-catenin accumulation as a predictor of hematogenous metastasis in human colorectal cancer. *Oncology* **59**, 302-9 (2000).
17. Hugh, T.J. et al. beta-catenin expression in primary and metastatic colorectal carcinoma. *Int J Cancer* **82**, 504-11 (1999).
18. Chung, G.G. et al. Tissue microarray analysis of beta-catenin in colorectal cancer shows nuclear phospho-beta-catenin is associated with a better prognosis. *Clin Cancer Res* **7**, 4013-20 (2001).
19. Lin, S.Y. et al. Nuclear localization of EGF receptor and its potential new role as a transcription factor. *Nat Cell Biol* **3**, 802-8 (2001).

Acknowledgements:

We thank Thomas D'Aquila, Mary Helie, Lori Charette, Diana Fischer, Eric Rimm and Paul Lizardi, for their help in this effort. We thank Jose Costa, Vince Marchesi, Al Reynolds, Richard Levenson, and Eric Fearon for review of the manuscript. This work was supported by grants from the William and Catherine Weldon Donaghue Foundation for Medical Research and grants from the NIH including: K0-8 ES11571, NIEHS (RLC), RO-1 GM57604 NCI (DLR) and US Army DAMD grant 01-000436.

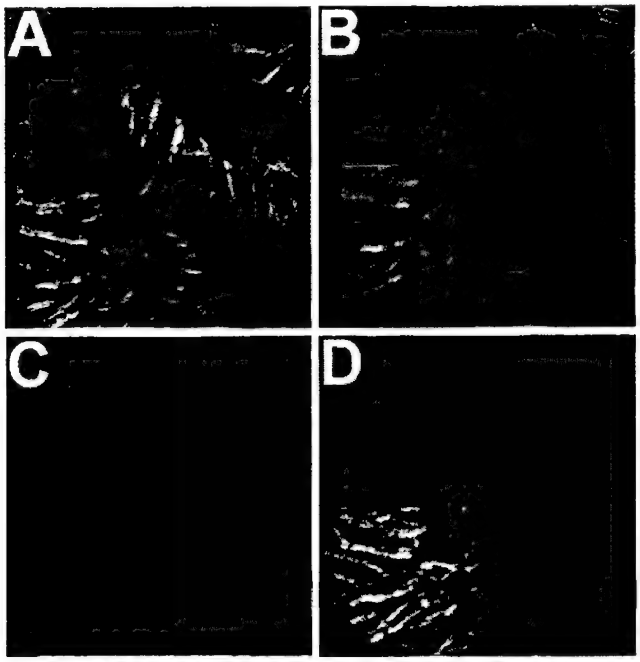

Figure 1. Rapid exponential image subtraction (RESA) allows the accurate assignment of sub-cellular compartments, and localization of a target antigen. Panel A: A pseudo-three color image of a colon carcinoma shows a significant degree of overlap between sub-cellular compartments: Blue = nuclei (DAPI), Green = tumor mask (binary alpha-catenin mask), Red = tumor cell membranes (RESA-processed alpha-catenin signal). **Panel B:** The signal intensity of a target antigen, _-catenin (inset), is redistributed according to the relative signal intensity of the compartments identified in panel A: Blue = nuclear-localized, Red = membrane-localized, Green = cytoplasmic. Note that the _-catenin expression in this tumor is predominantly membrane-associated, yet there is significant incorrectly assigned signal in the nucleus: magenta and blue pixels. **Panel C:** The compartment-specific signals in panel A are reassigned using the RESA algorithm, reducing the amount of overlapping signal by exponentially subtracting pixel intensity from an out-of-focus image. **Panel D:** The signal intensity from an exponentially subtracted image of the target antigen, _-catenin (inset) is then redistributed according to the compartments defined in panel C. This results in more accurate assignment of the target antigen to the membrane compartment (red pixels) with little expression in the nuclear compartment (blue pixels).

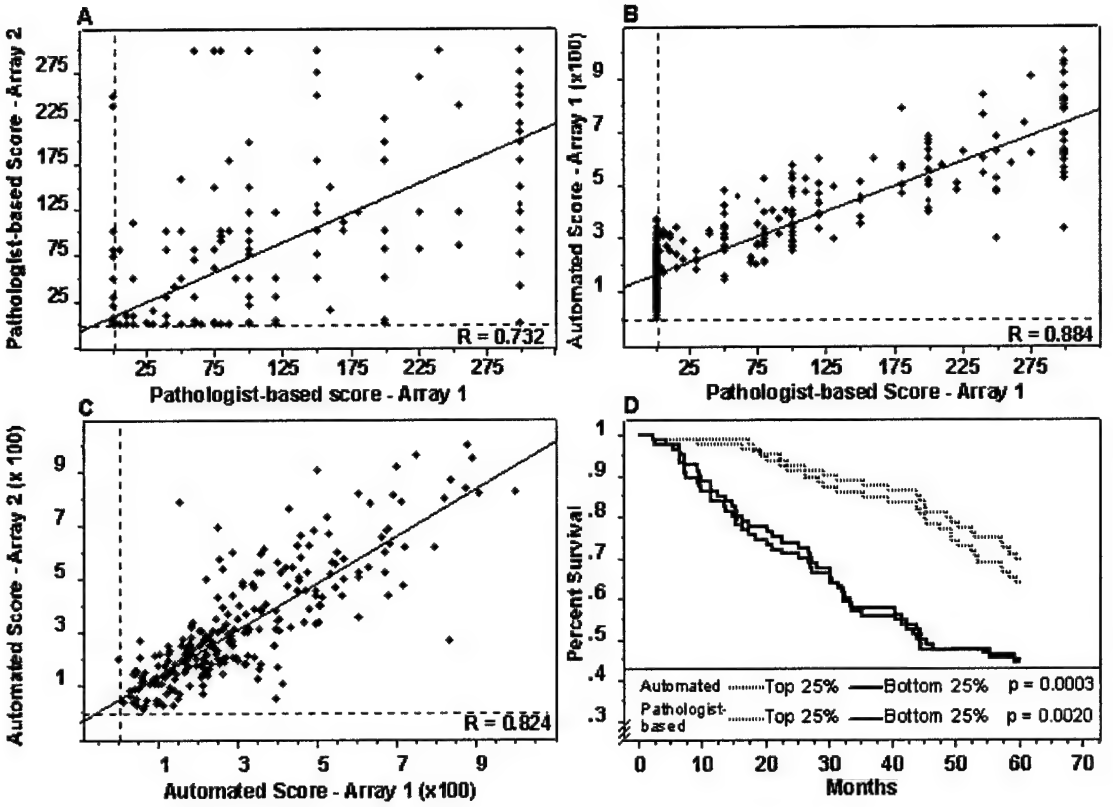

Figure 2. Automated and pathologist-based scoring of estrogen receptor shows a high degree of correlation and equal power in predicting outcome. **Panel A:** Two replicate tissue microarrays of 345 node-positive breast carcinomas were stained and scored by a pathologist using a traditional H-score algorithm. Note that despite good correlation (R = 0.732), there is significant discontinuity in the scoring pattern with numerous scores clustered around the high and low-end of the scale. **Panel B:** Automated analysis was performed using the RESA and PLACE algorithms and compared to the pathologist' score with a high degree of correlation (R = 0.884). **Panel C:** A comparison of two automated reads of two different microarrays derived from the same cohort also shows excellent correlation (R = 0.824). **Panel D:** To determine the potential of each system to predict outcome, survival curves were plotted for the top and bottom 25% ER expressers using each scoring system: Red = automated, Blue = pathologist-based. Both systems provided a statistically significant survival difference (P = 0.0003 and P = 0.0020, automated and pathologist-based, respectively).

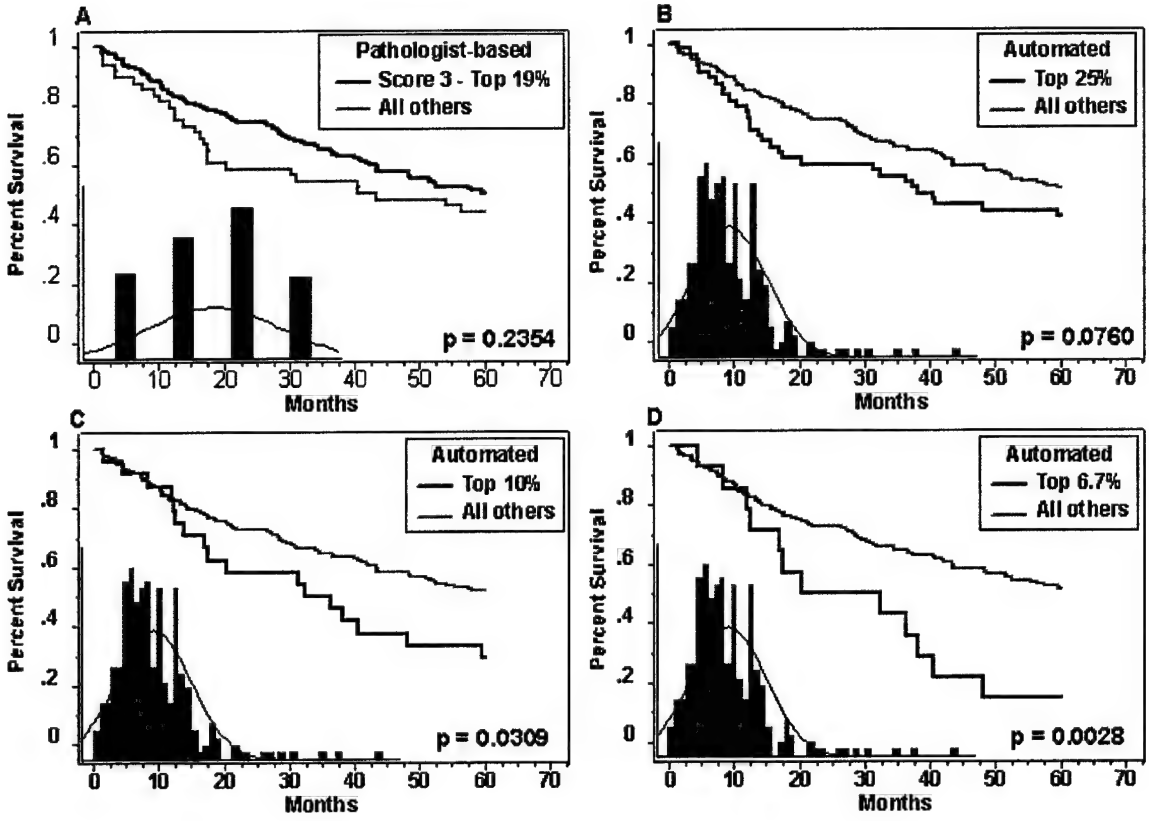

Figure 3. Increasing levels of nuclear beta-catenin associate with an increasingly poor prognosis. The relative amount of nuclear-associated beta-catenin was analyzed in our cohort of 310 colon carcinomas. **Panel A:** Analysis of nuclear beta-catenin levels in a pathologist-based analysis using a 4-point scale (0-3+) fails to find a significant survival difference when comparing tumors with the highest levels of nuclear localized _-catenin (3+, 19% of cases) versus the remaining cases (P = 0.2354). **Panel B through D:** In contrast, automated analysis of tumor subsets with higher and higher levels _-catenin show increasingly poorer survival, with increasing significance: *Panel B*: top 25%, P = 0.0760; *Panel C*: top 10%, P = 0.0309; *Panel D*: top 6.7%, P = 0.0028). Insets show the frequency distribution of intensity scores for each analysis with the selected subset in black (the x-axes for the insets are not shown but extend linearly from 1 to 1000). The relative risk of death at each level similarly increased with higher levels of nuclear _-catenin (RR = 1.404 {P = 0.0776, 95% C.I. 0.963 – 2.045} , RR = 1.740 {P = 0.0332, 95% C.I. 1.045 – 2.898}, and RR = 2.415 {P = 0.0038, 95% C.I. 1.330-4.386} respectively).

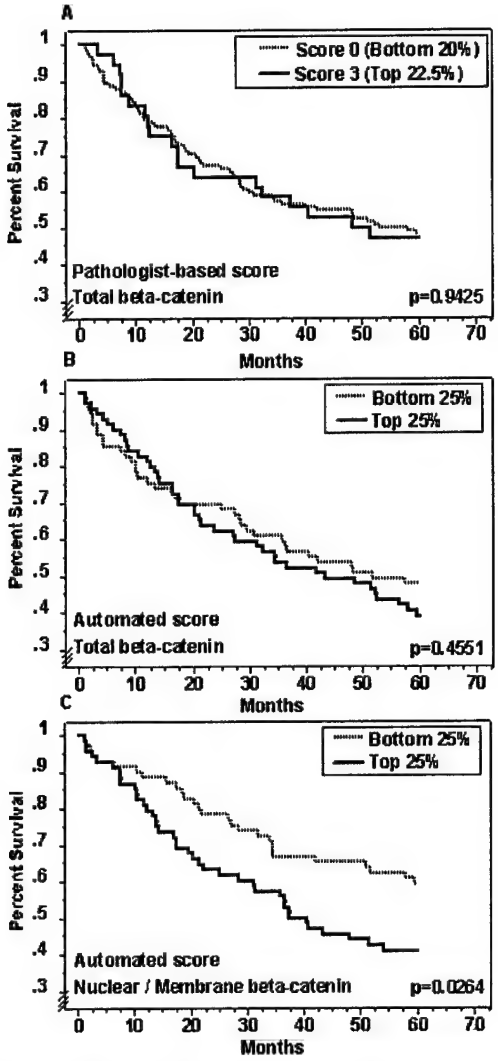

Figure 4. Unlike analyses of overall β-catenin expression, automated, subcellular localization of β-catenin can predict outcome in colon carcinoma. Panel A: Overall levels of beta-catenin expression in a cohort of 310 colon cancers were evaluated by a pathologist using a nominal 4-point scale (0-3+), with 20% of the tumors scored as 0 and 22.5% of the tumors scored as 3+. Comparison of tumors with the lowest (0) versus the highest (3+) overall _-catenin expression failed to find a significant survival difference (P = 0.9425). **Panel B:** Similarly, automated analysis of overall _-catenin levels, comparing cases in the top and bottom 25%, also failed to detect a survival difference (P = 0.4551). **Panel C:** In contrast, the ratio of nuclear to membrane _-catenin, as assessed by automated analysis demonstrates that tumors with higher relative nuclear expression do worse than those with higher relative membrane expression (P = 0.0264, top versus bottom 25% of cases), with a calculated relative risk of 1.718 (95% CI 1.059-2.787, P = 0.0284). Note that the nuclear/membrane ratio identifies a relatively large subset (25%) of tumors with poor prognosis, the majority of which are not identified by analyzing individual subcellular compartments. Indeed, comparison the tumors with the highest nuclear/membrane ratio (top 25%) versus the highest overall nuclear levels of _-catenin (top 25%), shows that there is only 47% overlap between the two subsets.

Tissue Microarray based studies of node-negative breast cancer patients show Met expression associated with worse outcome but not correlated with EGF family receptors.

Idris Tolgay Ocal, M.D., Marisa Dolled-Filhart, Thomas G. D'Aquila,
Robert L. Camp, M.D., Ph.D. and David L. Rimm, M.D., Ph.D.

Department of Pathology, Yale University School of Medicine, New Haven, Connecticut 06520

Running Title: *Met in Breast Cancer*

*Corresponding Author:
David L. Rimm M.D.-Ph.D.
Dept. of Pathology
Yale University School of Medicine
310 Cedar St.
New Haven, CT 06510

Phone 203-737-4204
FAX 203-737-5089
Email david.rimm@yale.edu

Keywords; HGF, c-met, adhesion, scatter factor, prognostic, Her2

Grant Support: Dr. Rimm is supported by a grant from the William and Catherine Weldon Donaghue Foundation for Medical Research, US Army Breast Cancer Research Program grant DAMD-17-01-0463 and grants from the NIH.

Condensed Abstract:
Tissue microarray technology is a new technology that provides an excellent method for comparison of expression of multiple markers. Here we examine a series of receptor tyrosine kinases and find that Met and FGFR are related, but Her2 and EGFR are not related to either, although they are related to each other. Only Met predicts poor outcome in our node negative cohort.

Abstract

Background:
Receptor tyrosine kinases (RTKs) have been shown to predict outcome in breast cancer. Although RTKs are a large family, HER2, EGFR, Met and others have all be shown to predict outcome. However, it is not clear that they are defining the same subpopulation of breast cancer cases. In this study we attempt to look at the relationship between RTKs on the basis of their ability to stratify a population on the basis of outcome.

Methods:
We used tissue microarray technology to examine 324 node-negative breast carcinoma cases with 20-40 year follow-up. Expression was assessed using immunohistochemical stains for Met, EFGR, FGFR, and Her2. Expression levels were assessed by 2 observers and relationships were analyzed. Standard pathology information, including tumor size, nuclear grade, tumor size, Ki67 and estrogen and progesterone receptor expressions was also collected.

Results:
RTK expression in our cohort reveals two strong relationships. Specifically, Her2 and EGFR showed similar expression patterns ($p<0.0001$) and Met cytoplasmic domain and FGFR-cytoplasmic staining showed similar expression patterns ($p<0.0001$) but no relationship was found between the two groups. Of these RTKs, only high levels of Met cytoplasmic domain showed significance as a prognostic marker defining a shortened survival compared to the rest of the population ($p=0.0035$, $RR=2.04$). In the same group of patients, Her2, hormone receptor status and other RTK family receptors did not correlate with outcome. In multivariate analysis only Met cytoplasmic domain and tumor size showed independent predictive value.

Conclusions:
In conclusion, we show that the cytoplasmic domain of Met shows a unique staining pattern and defines a set of patients unique from that described by over-expression of Her2, EGFR or hormone receptors. Furthermore, this group is tightly and independently associated with worse outcome.

Introduction

The single best indicator of disease free survival and overall survival in breast carcinoma is lymph node status [1]. Breast carcinomas with axillary lymph node metastases carry a 10 year recurrence rate approaching 70% [2]. However, in node-negative breast cancer, there are no standard markers used to predict outcome. In this group, predicting a worse outcome is of great importance because as many as 20% of node-negative women will eventually die of metastatic disease [3]. Numerous prognostic markers have been tested to try to predict outcome in this group. In large studies with long follow-up, the best predictors of outcome are tumor size, tumor grade, cathepsin-D, Ki-67, S-phase fraction, mitotic index, and vascular invasion. However, for various reasons, only tumor size and tumor grade enjoy broad acceptance for this patient group [4].

Tyrosine kinase receptors (RTK) are gaining attention as prognostic markers and as possible future predictive markers as the number of trials grows for bio-specific inhibitors. Among these, Her2 (erbB2, neu), EGFR and Met (c-met) have been documented as prognostically significant markers for invasive breast carcinomas predicting a worse prognosis. Her2 has been associated with outcome in node-positive tumors, but has not proven valuable in node-negative tumors [5]. Although some studies have suggested that EGFR is valuable [6], EGFR has been examined in over 25 studies, of which only about half suggest that over-expression is associated with poor outcome [7]. Met has also shown mixed results. Although we and others have found it useful in predicting worse outcome [8-10], others have not seen the relationship or found an opposite one [11]. Finally FGFR is less well studied, and no definitive prognostic value has been shown[12].

Slide to slide standardization is a longstanding problem in immunohistochemistry studies and may be one explanation for the variability seen in the studies above. Tissue microarray technology can eliminate this problem[13, 14]. Tissue microarrays are a method of placing very small samples of tissue from hundreds or thousands of patients on a single slide[15, 16]. The technology has now been used extensively and is the subject of multiple reviews [14, 17, 18] Tissue microarrays are especially well suited to comparison of expression between multiple prognostic markers. Here we revisit the issue of the prognostic value of 4 RTKs in node-negative breast cancer with an emphasis on the relationship between the markers. We studied expression patterns of Met, EGFR, FGFR, and HER2/neu on a cohort of patients with long term follow-up.

Material and Methods

Tissue Microarray Construction

The tissue microarrays were constructed as previously described [16], and reviewed recently[14]. Briefly, formalin fixed paraffin embedded tissue blocks containing breast cancer were retrieved from the archives of the Yale University Department of Pathology. Areas of invasive carcinoma were identified on corresponding H&E-stained slides, and the tissue blocks were cored and transferred to a recipient "master" block using a Tissue Microarrayer (Beecher Instruments, Madison, WI). Each core is 0.6mm in diameter and

spaced 0.8mm apart. After cutting the recipient block and transfer with an adhesive tape to coated slides for subsequent UV cross-linkage (Instrumedics, Inc., Hackensack, NJ), the slides were dipped in a layer of paraffin to prevent oxidation. The cohort node-negative breast cancer array was constructed from paraffin-embedded formalin-fixed tissue blocks from the Yale University Department of Pathology archives. The specimens were resected between 1962 and 1980, with a follow-up range between 4 months and 53.8 years, with a mean follow-up time of 15.6 years.

Grading

Nuclear grade was evaluated by one observer (ITO) only according to the following criteria: small uniform nuclei, no nucleoli, grade 1; moderate pleomorphism, prominent nucleoli, grade 2; marked pleomorphism, angulated nuclei, grade 3. Presence or absence of necrosis, mitotic count or histologic parameters could not be included in the grading criteria, because of the small size of the area evaluated. Eighty cases (26%) out of 306 included were grade I, 170 (56%) were grade II and 56 (18%) were scored grade III.

Tumor Size

Information about the tumor size was obtained from the gross descriptions in the original pathology reports. For staging for statistical analyses, a tumor size greater than 2.0cm was considered "large" and others "small".

Immunohistochemistry

The tissue microarray slides were deparaffinized with xylene rinses and then transferred through two changes of 100% ethanol. Endogenous peroxidase activity was blocked by a 30-minute incubation in a 2.5% hydrogen peroxide/methanol buffer. Antigen retrieval was performed by boiling the slides in a pressure cooker filled with a sodium citrate buffer (pH 6.0). After antigen retrieval, the slides were incubated with 0.3%BSA/1XTBS for one hour at room temperature in order to reduce non-specific background staining, followed by a series of 2 minute rinses in 1XTBS, TBS/0.01% Triton, 1XTBS. Primary antibody was applied for one hour at room temperature. Dilutions for the RTKs were Met, 1:1 (note this antibody was provided as a culture supernate from Zymed), EGFR 1:200, FGFR 1:300 and Her2 1:8000. After a series of TBS rinses as described above, bound antibody was detected by using an anti-rabbit horseradish peroxidase-labeled polymer secondary antibody from the DAKO Envision TM + System (DAKO, Carpinteria, CA). The slides were rinsed in the TBS series, visualized with a 10 minute incubation of liquid 3,3'-diaminobenzidine in buffered substrate (DAKO, Carpinteria, CA) for 10 minutes. Finally, the slides were counterstained with hematoxylin, and mounted with Immunomount (Shandon, Pittsburgh, PA). Immunohistochemical staining was also done for estrogen receptor, progesterone receptor and Her2 as described previously [19]. Ki-67 expression was assessed using purified anti-human mAb (1:200, overnight incubation, Pharmingen, San Diego, CA). The Met antibody 3D4, was provided as part of a collaborative arrangement with Zymed Laboratories Inc (South San Francisco, CA), the EGFR antibody was EGFR 1005 (Cat.# SC03) from Santa Cruz Biotechnology Inc, (Santa Cruz, CA) and the FGFR1 antibody was FLG C15 (Cat #SC121) from Santa Cruz Biotechnology Inc

Evaluation of Immunohistochemical staining

For each spot, the regions of most intense and/or predominant staining pattern were scored by eye. Traditionally, immunohistochemistry scoring of stain intensity includes a variable for the area percentage stained with the specimen, but due to the small size of the spot (0.6mm in diameter) and the fact that the spots are often homogenous, no area variable is included. The nuclear and/or cytoplasmic (ER/PR, EGFR, FGFR, Ki67 and Met cytoplasmic domains) staining was determined separately for each specimen. The staining intensity was graded on the following scale: 0, no staining; 1, weak staining; 2, moderate staining; 3, intense staining. The membranous staining was determined for HER2/neu and EGFR. The staining intensity for membranous staining was graded on the following scale: 0, no staining; 1, incomplete staining; 2, weak but complete staining plasma membrane encircling the entire cell; 3, intense complete staining. Again, we did not take the percentage of cells with staining in consideration, because of the small size of the tumor sections. For specimens that were uninterpretable, a score of N/A was given. Only FGFR showed distinctly separate nuclear and cytoplasmic staining characteristics and for this antibody, nuclear and cytoplasmic staining were scored individually. Scoring of the tissue microarrays was completed by two independent observers (ITO and MDF) for EGFR, FGFR-nuclear staining, FGFR-cytoplasmic staining and Met with very high correlation between scorers ($p<0.0001$). For the antibodies with established staining characteristics in literature (ER, PR, Her2 and Ki-67), scoring was performed by one observer (ITO or MDF). Ki-67 was considered positive if more than 10% of the nuclei were stained. Frequency distributions for these markers were in the range of other works seen in the literature, validating this cohort (see table 1).

Statistical Analysis

All analyses were completed using Statview 5.0.1 software (SAS Institute Inc., Cary, NC). The correlation between the scores of both scorers and the relationships among the different immunohistochemical and clinicopathological parameters were measured using the Chi Squared test. The prognostic significance of the parameters on overall survival was calculated by multivariate analysis using the Cox proportional hazards model. Survival curves were calculated using the Kaplan-Meier survival analysis method with the differences estimated using the Mantel-Cox log-rank test.

Results

The RTKs analyzed in this study showed a variety of staining patterns that are summarized in figure 1. Although RTKs are all present at the membrane, there are now numerous studies that show they can be found in other locations in the cell including both the cytoplasm and more recently, the nucleus[20]. The patterns we saw included both conventional membranous patterns as well as both cytoplasmic and nuclear staining (described in detail ion the methods section). The pattern of expression of each RTK was determined after examining the entire array to determine what patterns were most prevalent. The question of how to divide subjective ordinal staining patterns is always controversial. Here we tried to use breakpoints suggested previously in the literature when they were present [8, 10], but in some cases (FGFR) there were not clear precedents in

the literature or the literature was inconsistent (EGFR). A summary of the expression pattern of each RTK is shown in table 2 with the definition of the cut-points.

In this study, we were primarily concerned with the relationship between each RTK and the relationship of each RTK to known prognostic markers. We calculated Chi square P values to determine the relationship among expression patterns of individual parameters. Table 3 shows the Chi square p-values for all the parameters studied in node negative breast carcinomas. Of particular note are the highly significant correlations in immunohistochemical expression patterns of Her2 and EGFR, and Met and FGFR-cytoplasmic staining ($p<0.0001$). ER expression correlated highly with PR and they both showed significant inverse relationship to Her2 expression.

Survival analyses. Cox univariate analyses at 10 years for all the variables studied in node negative breast carcinomas are shown in Table 3. As expected, Ki67 and tumor size correlated with poorer survival (P=0.04, RR=1.648 and P=0.0008, RR=2.201 respectively). Among the other variables studied, only cytoplasmic domain of Met showed statistically significant correlation with a worse prognosis and shortened survival (P=0.0035) with a relative risk of 2.041. Survival curves were produced for each variable, but only Met reached statistical significance using the Mantel-Cox log rank test (Figure 2). To determine the independent predictive value of Met expression a multivariate analysis was done using the Cox Proportional hazards model. In multivariate analysis, Met retained its significance as a predictor of worse outcome even when the model contained all of the conventional prognostic variables as well as all other RTKs tested (RR=1.86, p=0.011).

Discussion

In node negative breast carcinomas we are still unable to discern the 15-20% of patients that will eventually succumb to their disease [3, 21]. A long list of potential molecular markers of poorer prognosis have been suggested in this group of breast carcinoma including Ki-67[22, 23], cathepsin-D[24], Her2[22, 25-27], p5[26, 28], low levels of nm23[29], and hormone receptors[23, 26, 30]. There are also reports that histologic findings such as tumor grade, size, mitotic index, vascular invasion, and cytometric data such as DNA ploidy and S-phase fraction are helpful[22, 23, 26, 30-34]. To date, there are no universally accepted markers for this group.

We evaluated our tissue microarray cohort using this traditional markers for breast cancer. Hormone receptors showed similar staining patterns and frequencies compared to the widely reported values in breast carcinoma. ER was negative in 35% of cases and PR was negative in 40%. Neither ER nor PR expression was predictive of survival in our group. This finding is consistent with many other published studies that show limited or no value for ER and PR in node negative breast carcinomas as prognostic markers[30, 35]. Her2 has been proven to predict a worse prognosis in invasive, node positive breast lesions and is a better predictor of survival than hormone receptors in this group[36]. However, its predictive value in node-negative lesions is still controversial with

conflicting data in the literature [22, 25-27]. Our findings seem to support the hypothesis that Her2 is not a reliable prognostic marker in this group of patients.

EGFR expression in breast carcinomas have been reported in both node negative and node positive breast carcinoma as associated with poor prognosis; however, the clinical significance and association with disease free (DFS) and overall survival (OS) statistics show mixed results in different studies [37-39]. Tsutsui reported that EGFR carries a prognostic significance only for DFS in node negative tumors on multivariate analysis while in Torregrosa's series EGFR failed to show a statistical significance as a prognostic marker in node-negative tumors, but was associated with worse prognosis influencing DFS in node positive patients. On the other hand, Seshadri concluded that the expression of EGFR was not a predictor of poor prognosis in node negative breast carcinoma patients at all. Our findings did not show any predictive value of EGFR as a marker of prognosis in node negative patients.

Our group has been particularly interested in the HGF receptor, Met (or C-met). Met is a dimeric tyrosine kinase growth factor receptor whose activation is associated with increased invasion, motogenesis and morphogenesis [40]. In the breast, Met is expressed in normal ductal and lobular epithelium and functions in both the embryonic development and subsequent remodeling of the breast [41].

Our previous studies have shown that the expression of Met receptor in patients with invasive breast carcinoma is of significant predictive value in determining patient survival even in node negative patients [8, 9]. The Van der Woude group has also found this relationship between Met expression and outcome [10], but we and others have found no relationship in some cohorts (unpublished data). We believe this variability is due to antibody selection. Many studies have used antibodies to the c-terminal domain produced by Santa Cruz Biotechnology Inc using a peptide coding for the C-terminal 28 amino acids. We have found significant lot to lot variability with this antibody. Recently, Zymed Laboratories Inc and others have produced monoclonal antibodies to the C-terminus. In another study from our lab that was concurrently submitted (Kang et al), we compared the results with this antibody and another monoclonal to the extracellular domain of Met. We showed similar results with the cytoplasmic domain (to those reported here) but the extracellular domain is very different. Although there is high correlation of expression, over-expression as assessed by the antibody to the cytoplasmic domain, selects a group of patients with worse outcome, while the extracellular domain antibody does not (Kang et al, manuscript submitted). We believe this may be a function of either cleavage or activation of Met. It is notable that the other study that found a relationship between Met expression and outcome also used a monoclonal antibody (generated by the Van der Woude lab) [10] and that our own previous studies used a polyclonal antibody made to a cytoplasmic domain peptide[8].

The finding that Met overexpression predicts poor outcome raises the question of its relationship to other RTKs, especially EGFR and Her2 which have also been implicated as prognostic variables. This work shows that the tumors that overexpress Met are unique from those that express Her2 and EGFR. EGFR and Her2 are closely related

members of the erbB oncogene family, so it is not surprising that there is a high correlation between expression of these two proteins. Met is not a member of this family and thus its over-expression appears to be unrelated to erbB family RTKs.

Even though there is not a close relationship between Met and erbB family RTKs, there is a tight direct correlation between Met expression and the cytoplasmic staining pattern seen with the FGFR antibody ($P<0.0001$). To our knowledge, the relationship between FGFR and Met has never been reported before in breast carcinoma. Coordinated actions of several growth factors and their receptors including FGFR and C-met have been reported in hepatic lesions [42] and normal morphogenesis in the uterus [43]; however, the molecular basis of such an interaction, if present, is still unknown. Although FGFR is a membrane-based tyrosine kinase receptor, immunohistochemical evaluation for FGFR has not been standardized as a marker for immunohistochemistry. Thus we scored cytoplasmic and nuclear staining for FGFR separately. Although neither cytoplasmic nor nuclear staining for FGFR was a statistically significant prognostic marker in our study, there was a high correlation between Met expression and FGFR cytoplasmic expression.

RTKs recently have gained great interest as both markers of prognosis and also as promising targets for novel chemotherapeutic options. The HER2/ herceptin pair is the first FDA approved example of numerous RTK related therapeutics currently in clinical trials. This interest raises the question of the relationships between different RTKs. Although data on the cross reactivity of kinase-based therapeutics is not yet widely available, there are many of types of RTK receptors over-expressed in breast cancers.

In summary, the current study indicates that the expression of Met, as assessed using a cytoplasmic domain monoclonal antibody, in patients with node-negative invasive breast carcinoma is of significant predictive value in determining patient survival. Its predictive value exceeds all conventional prognostic markers assessed in this cohort as well as the other RTKs tested. The group of patients that over-express this protein are unrelated to the group that overexpresses either EGFR or Her2. Thus, in the future, determination of prognosis in node-negative invasive breast cancers may be improved by assessment of the level of expression of Met using a cytoplasmic domain monoclonal antibody.

References:

1. Fitzgibbons PL, Page DL, Weaver D, Thor AD, Allred DC, Clark GM, et al. Prognostic factors in breast cancer. College of American Pathologists Consensus Statement 1999. *Arch Pathol Lab Med* 2000; 124:966-78.
2. Elledge RM, McGuire WL. Prognostic factors and therapeutic decisions in axillary node-negative breast cancer. *Annu Rev Med* 1993; 44:201-10.
3. Fisher ER, Costantino J, Fisher B, Redmond C. Pathologic findings from the National Surgical Adjuvant Breast Project (Protocol 4). Discriminants for 15-year survival. National Surgical Adjuvant Breast and Bowel Project Investigators. *Cancer* 1993; 71:2141-50.
4. Mirza AN, Mirza NQ, Vlastos G, Singletary SE. Prognostic factors in node-negative breast cancer: a review of studies with sample size more than 200 and follow-up more than 5 years. *Ann Surg* 2002; 235:10-26.
5. Lohrisch C, Piccart M. HER2/neu as a predictive factor in breast cancer. *Clin Breast Cancer* 2001; 2:129-35; 136-7.
6. Pawlowski V, Revillion F, Hebbar M, Hornez L, Peyrat JP. Prognostic value of the type I growth factor receptors in a large series of human primary breast cancers quantified with a real-time reverse transcription-polymerase chain reaction assay. *Clin Cancer Res* 2000; 6:4217-25.
7. Nicholson RI, Gee JM, Harper ME. EGFR and cancer prognosis. *Eur J Cancer* 2001; 37 Suppl 4:S9-15.
8. Camp RL, Rimm EB, Rimm DL. Met expression is associated with poor outcome in patients with axillary lymph node negative breast carcinoma. *Cancer* 1999; 86:2259-65.
9. Ghoussoub RA, Dillon DA, D'Aquila T, Rimm EB, Fearon ER, Rimm DL. Expression of c-met is a strong independent prognostic factor in breast carcinoma. *Cancer* 1998; 82:1513-20.
10. Tsarfaty I, Alvord WG, Resau JH, Altstock RT, Lidereau R, Bieche I, et al. Alteration of Met protooncogene product expression and prognosis in breast carcinomas. *Anal Quant Cytol Histol* 1999; 21:397-408.
11. Nakopoulou L, Gakiopoulou H, Keramopoulos A, Giannopoulou I, Athanassiadou P, Mavrommatis J, et al. c-met tyrosine kinase receptor expression is associated with abnormal beta-catenin expression and favourable prognostic factors in invasive breast carcinoma. *Histopathology* 2000; 36:313-25.
12. Dickson C, Spencer-Dene B, Dillon C, Fantl V. Tyrosine kinase signalling in breast cancer: fibroblast growth factors and their receptors. *Breast Cancer Res* 2000; 2:191-6.
13. Torhorst J, Bucher C, Kononen J, Haas P, Zuber M, Kochli OR, et al. Tissue microarrays for rapid linking of molecular changes to clinical endpoints. *Am J Pathol* 2001; 159:2249-56.
14. Rimm DL, Camp RL, Charette LA, Costa J, Olsen DA, Reiss M. Tissue microarray: a new technology for amplification of tissue resources. *Cancer J* 2001; 7:24-31.
15. Wan WH, Fortuna MB, Furmanski P. A rapid and efficient method for testing immunohistochemical reactivity of monoclonal antibodies against multiple tissue samples simultaneously. *J Immunol Methods* 1987; 103:121-9.

16. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 1998; 4:844-7.
17. Bubendorf L, Nocito A, Moch H, Sauter G. Tissue microarray (TMA) technology: miniaturized pathology archives for high-throughput in situ studies. *J Pathol* 2001; 195:72-9.
18. Skacel M, Skilton B, Pettay JD, Tubbs RR. Tissue microarrays: a powerful tool for high-throughput analysis of clinical specimens: a review of the method with validation data. *Appl Immunohistochem Mol Morphol* 2002; 10:1-6.
19. Camp RL, Charette LA, Rimm DL. Validation of tissue microarray technology in breast carcinoma. *Lab Invest* 2000; 80:1943-9.
20. Lin SY, Makino K, Xia W, Matin A, Wen Y, Kwong KY, et al. Nuclear localization of EGF receptor and its potential new role as a transcription factor. *Nat Cell Biol* 2001; 3:802-8.
21. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer* 1989; 63:181-7.
22. Rudolph P, Olsson H, Bonatz G, Ratjen V, Bolte H, Baldetorp B, et al. Correlation between p53, c-erbB-2, and topoisomerase II alpha expression, DNA ploidy, hormonal receptor status and proliferation in 356 node-negative breast carcinomas: prognostic implications. *J Pathol* 1999; 187:207-16.
23. Seshadri R, Leong AS, McCaul K, Firgaira FA, Setlur V, Horsfall DJ. Relationship between p53 gene abnormalities and other tumour characteristics in breast-cancer prognosis. *Int J Cancer* 1996; 69:135-41.
24. Foekens JA, Look MP, Bolt-de Vries J, Meijer-van Gelder ME, van Putten WL, Klijn JG. Cathepsin-D in primary breast cancer: prognostic evaluation involving 2810 patients. *Br J Cancer* 1999; 79:300-7.
25. Sjogren S, Inganas M, Lindgren A, Holmberg L, Bergh J. Prognostic and predictive value of c-erbB-2 overexpression in primary breast cancer, alone and in combination with other prognostic markers. *J Clin Oncol* 1998; 16:462-9.
26. Reed W, Hannisdal E, Boehler PJ, Gundersen S, Host H, Marthin J. The prognostic value of p53 and c-erb B-2 immunostaining is overrated for patients with lymph node negative breast carcinoma: a multivariate analysis of prognostic factors in 613 patients with a follow-up of 14-30 years. *Cancer* 2000; 88:804-13.
27. Press MF, Bernstein L, Thomas PA, Meisner LF, Zhou JY, Ma Y, et al. HER-2/neu gene amplification characterized by fluorescence in situ hybridization: poor prognosis in node-negative breast carcinomas. *J Clin Oncol* 1997; 15:2894-904.
28. Silvestrini R, Benini E, Daidone MG, Veneroni S, Boracchi P, Cappelletti V, et al. p53 as an independent prognostic marker in lymph node-negative breast cancer patients. *J Natl Cancer Inst* 1993; 85:965-70.
29. Barnes R, Masood S, Barker E, Rosengard AM, Coggin DL, Crowell T, et al. Low nm23 protein expression in infiltrating ductal breast carcinomas correlates with reduced patient survival. *Am J Pathol* 1991; 139:245-50.
30. Fisher B, Redmond C, Fisher ER, Caplan R. Relative worth of estrogen or progesterone receptor and pathologic characteristics of differentiation as indicators of prognosis in node negative breast cancer patients: findings from National Surgical Adjuvant Breast and Bowel Project Protocol B-06. *J Clin Oncol* 1988; 6:1076-87.

31. Rosen PR, Groshen S, Saigo PE, Kinne DW, Hellman S. A long-term follow-up study of survival in stage I (T1N0M0) and stage II (T1N1M0) breast carcinoma. *J Clin Oncol* 1989; 7:355-66.
32. Railo M, Lundin J, Haglund C, von Smitten K, von Boguslawsky K, Nordling S. Ki-67, p53, Er-receptors, ploidy and S-phase as prognostic factors in T1 node negative breast cancer. *Acta Oncol* 1997; 36:369-74.
33. Balslev I, Christensen IJ, Rasmussen BB, Larsen JK, Lykkesfeldt AE, Thorpe SM, et al. Flow cytometric DNA ploidy defines patients with poor prognosis in node-negative breast cancer. *Int J Cancer* 1994; 56:16-25.
34. Aaltomaa S, Lipponen P, Eskelinen M, Kosma VM, Marin S, Alhava E, et al. Mitotic indexes as prognostic predictors in female breast cancer. *J Cancer Res Clin Oncol* 1992; 118:75-81.
35. Fisher ER, Redmond C, Fisher B. Prognostic factors in NSABP studies of women with node-negative breast cancer. National Surgical Adjuvant Breast and Bowel Project. *J Natl Cancer Inst Monogr* 1992:151-8.
36. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 1987; 235:177-82.
37. Torregrosa D, Bolufer P, Lluch A, Lopez JA, Barragan E, Ruiz A, et al. Prognostic significance of c-erbB-2/neu amplification and epidermal growth factor receptor (EGFR) in primary breast cancer and their relation to estradiol receptor (ER) status. *Clin Chim Acta* 1997; 262:99-119.
38. Tsutsui S, Ohno S, Murakami S, Hachitanda Y, Oda S. Prognostic value of epidermal growth factor receptor (EGFR) and its relationship to the estrogen receptor status in 1029 patients with breast cancer. *Breast Cancer Res Treat* 2002; 71:67-75.
39. Seshadri R, McLeay WR, Horsfall DJ, McCaul K. Prospective study of the prognostic significance of epidermal growth factor receptor in primary breast cancer. *Int J Cancer* 1996; 69:23-7.
40. Comoglio PM, Boccaccio C. Scatter factors and invasive growth. *Seminars in Cancer Biology* 2001; 11:153-165.
41. Niranjan B, Buluwela L, Yant J, Perusinghe N, Atherton A, Phippard D, et al. HGF/SF: a potent cytokine for mammary growth, morphogenesis and development. *Development* 1995; 121:2897-908.
42. Hu Z, Evarts RP, Fujio K, Omori N, Omori M, Marsden ER, et al. Expression of transforming growth factor alpha/epidermal growth factor receptor, hepatocyte growth factor/c-met and acidic fibroblast growth factor/fibroblast growth factor receptors during hepatocarcinogenesis. *Carcinogenesis* 1996; 17:931-8.
43. Taylor KM, Chen C, Gray CA, Bazer FW, Spencer TE. Expression of messenger ribonucleic acids for fibroblast growth factors 7 and 10, hepatocyte growth factor, and insulin-like growth factors and their receptors in the neonatal ovine uterus. *Biol Reprod* 2001; 64:1236-46.

Table 1
Distribution of expression of standard prognostic markers in cohort on tissue microarray

	ER	PR	Her2	Ki67
Positive[1](%)	57	52	14	60
Negative (%)	43	48	86	40

1. The definitions for ordinal scoring of expression and the selection of the cut points to define positive staining for each marker are described in the methods and results sections.

Table 2
Distribution of expression of RTKs on tissue microarray

	Her2	EGFR	FGFR(nuc)	FGFR(ctyo)	Met (cyto)
Positive[1](%)	14	10.5	48	68	22
Negative (%)	86	89.5	52	32	78

1. The definitions for ordinal scoring of expression and the selection of the cut points to define positive staining for each marker are described in the methods and results sections.

Table 3
Chi Square analysis of the relationship between expression of each marker

	ER	**PR**	**HER2**	**EGFR**	**FGFR-c**	**FGFR-n**	**MET**	**KI-67**
ER	*							
PR	**<.0001**	*						
HER2	**0.0003†**	**0.0011†**	*					
EGFR	**0.0016†**	**0.0010†**	**<.0001**	*				
FGFR-c	0.17	0.24	0.096	0.35	*			
FGFR-n	0.90	0.17	0.074	0.11	**0.001**	*		
MET	**0.019†**	0.19	0.54	0.77	**<.0001**	0.12	*	
KI-67	0.55	**0.028†**	**0.038**	0.053	**0.007**	0.26	**0.03**	*
Nuc GRADE	**0.015†**	**0.017†**	**0.019**	**0.0475**	0.065	**0.0016**	**<.0001**	**<.0001**

Statistical analyses of correlation patterns among the parameters studied in node negative breast carcinomas. Statistically significant values shown in bold. n=324. † indictates an inverse correlation.

Table 4.
Univariate analysis of Conventional and RTK markers

	RR	P	95% LOWER	95% UPPER
ER	1.031	0.9042	0.632	1.681
PR	0.946	0.8184	0.587	1.522
HER2	1.118	0.7236	0.603	2.071
EGFR	1.029	0.9400	0.494	2.141
FGFR-c	1.078	0.7456	0.685	1.695
FGFR-n	1.338	0.2200	0.840	2.129
MET	**2.041**	**0.0035**	**1.264**	**3.295**
KI-67	**1.648**	**0.0418**	**1.019**	**2.665**
Nuc GRADE	1.033	0.9120	0.578	1.846
Tumor Size	**2.201**	**0.0008**	**1.391**	**3.481**

Statistically significant values shown in bold.

Figure 1:

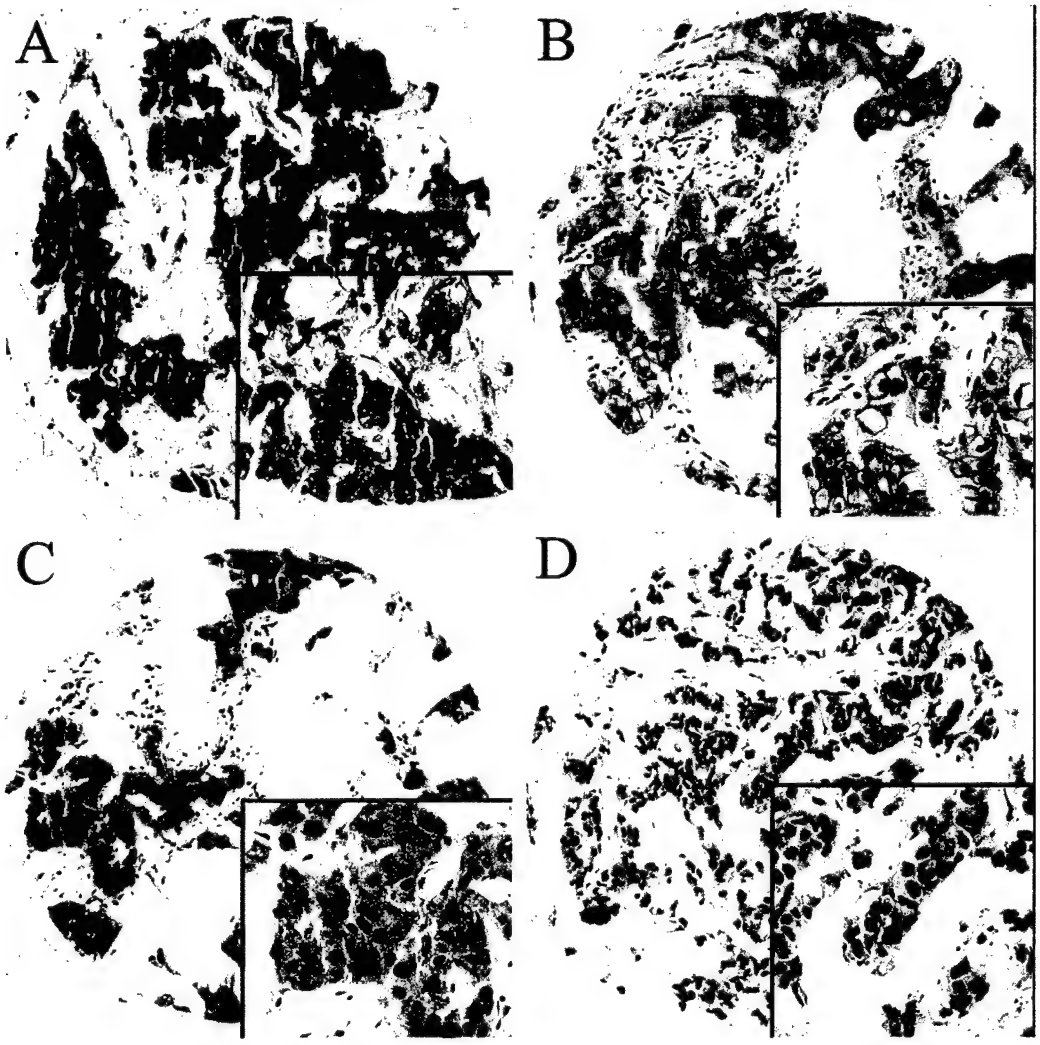

Figure 1: Examples of the tissue microarray immunostains for high level staining (3+) for A) Met, B) EGFR, C) FGFR cytoplasmic staining, and D) FGFR nuclear staining

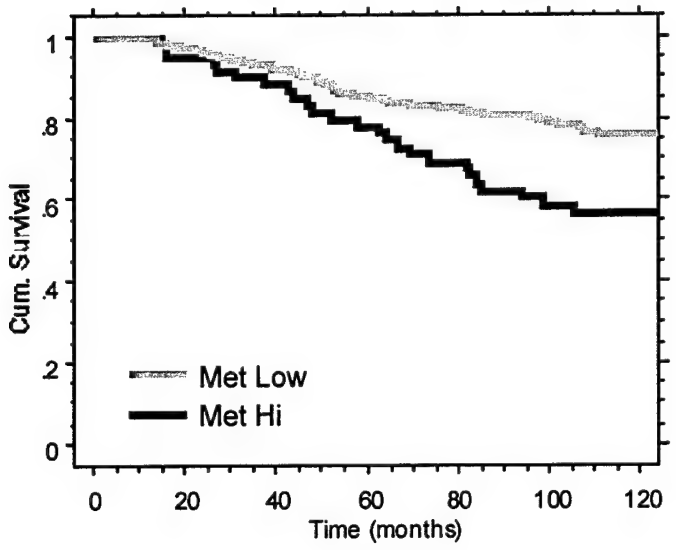

Figure 2. Kaplan-Meier survival analysis demonstrates that Met expression as assessed by antibodies to the cytoplasmic domain have significant predictive value for the survival of patients with breast carcinoma (P=0.0029, Mantel-Cox log-rank test). Time is indicated in months.

Quantitative analysis of breast cancer tissue microarrays shows both high and normal levels of HER2 expression are associated with poor outcome.

R.L. Camp, M. Dolled-Filhart, B. King, and D.L. Rimm

Departments of Pathology, Genetics, and Therapeutic Radiology, Yale University, School of Medicine

Summary

Over-expression of HER2/neu (c-erb B2) in breast cancer is associated with poor outcome particularly for node-positive patients. Clinically, HER2 expression is analyzed using either immunohistochemistry or fluorescence in situ hybridization (FISH) on patient samples. Protein level expression is determined manually on a four-point scale (0 to 3+). We have recently developed a system for the automatic and quantitative analysis of immunohistochemical scores using a continuous scale (AQUA). Using tissue microarrays, we compared analyses of HER2 on a cohort of 300 node-positive breast cancers using both manual and automated techniques. Results from both systems were highly correlated, and as expected, both systems identified a population (10-15%) of tumors expressing high levels of HER2 with poor 30-year disease related survival. Using AQUA analysis, we analyzed 84 samples of normal breast epithelium and found that they expressed a low but detectable level of Her-2. When we compared the HER2 expression of tumors in our cohort, we found that 17.5% exhibited HER2 levels in the range of normal epithelium. Survival analysis revealed that these tumors were as aggressive as HER2 overexpressing tumors. Furthermore, multivariate analysis revealed that both normal and high-HER2 expression retained independent prognostic value. Our studies suggest that in situ quantitative measurement of HER2 stratifies breast tumors into three expression levels: normal, intermediate, and high, where both normal and high levels are associated with a worse outcome.

Introduction

HER2 (neu or erb-B2) – a member of the epidermal growth factor family – is genetically amplified and overexpressed in aggressive breast cancers. High-levels of HER2 are associated with poor prognosis, particularly in node-positive breast carcinoma patients. Recently, a targeted therapeutic against HER2 has been developed. Trastuzumab (Herceptin) is a humanized monoclonal antibody directed against the extracellular domain of HER2. Treatment with Herceptin for patients with metastatic breast carcinoma has shown therapeutic benefit, especially when combined with conventional chemotherapeutic agents. The association between HER2 expression and Herceptin response has stimulated renewed interest in accurately assessing HER2 amplification and over-expression.

Overexpression of HER2 at the either the protein or amplification at the DNA level is assessed by immunohistochemical (IHC) staining or fluorescence *in situ* hybridzation

(FISH) of formalin-fixed tumor sections, respectively. Clinically, a pathologist interprets IHC and FISH results via manual examination. Although both techniques are used to stratify patients for Herceptin treatment, recent studies have demonstrated a significant lack of reproducibility between laboratories performing the tests and even between the different commercially available tests[1-3]. Causes for this deficiency involve both differences in sample preparation as well as differences in interpretation. In the case of HER2 IHC, samples are read on a nominal, four-point scale (0-3+). However, as with all biologic markers, actual HER2 expression is a continuous spectrum, not easily converted into a four-point scale. Even for the trained eye of a pathologist, accurate distinction between nominal categories (e.g. 2+ vs. 3+) is difficult and often arbitrary.

Toward the goal of automated analysis of protein expression on tissue microarrays, we have developed a system for the automated quantitative analysis of stained histologic sections (AQUA)[4]. As with an enzyme-linked immunoabsorbent assays (ELISA), AQUA provides a highly reproducible analysis of target signal expression using a continuous, rather than nominal scale. Previously, we have demonstrated how this system can accurately predict patient outcome for estrogen receptor expression in breast cancer, and beta-catenin expression in colon cancer[4]. The system has the added advantage of performing subcellular compartmentalization of target signals using a novel co-localization technique. To determine if the AQUA system would be beneficial for the analysis of HER2, we analyzed a cohort of 300 node-positive breast cancers along with 84 samples of normal mammary epithelium. As expected, we identified a population of HER2-high tumors with poor outcome. In addition we identified a second subset expressing HER2 levels similar to normal epithelium that also had a poor outcome. Two earlier studies using biochemical techniques (e.g. ELISA) for the analysis of HER2 protein expression suggested a bimodal distribution for HER2 with both low and high levels correlating with known markers of tumor aggression [5-7], however a third saw no such distribution [8]. Because ELISA cannot assess antigen levels in archival formalin-fixed tissue, is not generally amenable to studying large cohorts with long-term follow-up data. Alternatively, we used immunohistochemical stains on archival tissue to validate the bimodal distribution of HER2 by showing that tumors expressing both high and normal HER2 levels, exhibited poor 30-year disease-specific survival.

Materials and Methods

Tissue microarray design

Paraffin-embedded formalin-fixed specimens from 300 cases of node-positive invasive breast carcinoma were identified from the archives of the Yale University, Department of Pathology as available from 1962-1977, with a mean follow-up time of 9.6 years. No patients received Herceptin during the study period. Complete treatment information was unavailable for the entire cohort; however most patients were treated with local radiation, and approximately 15% were given chemotherapy consisting primarily of adriamycin, cytoxan, and 5-fluorouricil. Approximately 27% subsequently received tamoxifen (post-1978). Seven patients had biopsy proven stage IV disease at the time of diagnosis. In constructing the microarrays, areas of invasive carcinoma, away from *in situ* lesions and

normal epithelium, were identified and two cores 0.6 mm cores were taken. Each core was arrayed into recipient blocks in a 1mm-spaced grid covering approximately 1 square inch, and five-micron thick sections were cut and processed as previously described[9]. An additional microarray consisting of 84 samples of normal epithelium was also constructed from samples of normal ducts and lobules taken from breast cancer patients. Samples were taken away from areas of tumor and assessed histologically to ensure that they were unaffected by atypical hyperplasia or carcinoma in situ.

Immunohistochemistry

In brief, pre-cut paraffin-coated tissue microarray slides were deparaffinized, and subjected to pressure-cooking for antigen retrieval [10]. Slides were preincubated with 0.3% bovine serum albumin in 0.1M tris-buffered saline (pH 8.0) (BSA/TBS) for 30 min. at room temperature. For both manual and automated analysis, slides were incubated with polyclonal anti-HER2 (1:200, DAKO Corp., Carpinteria, CA) diluted in BSA/TBS for 1 hr at room temperature. Prior analysis of titrations of the HER2 antibody demonstrated that higher dilutions of anti-HER2 antibody (1:1000 ~ 1:8000) more accurately define the HER2-high from the HER2-intermediate populations, whereas lower dilutions (1:50 ~ 1:500) distinguish the HER2-normal from HER2-intermediate populations (unpublished observations). In this study we used concentration (1:200) that sufficiently distinguishes all three populations. Slides were washed 3x 5 min with TBS containing 0.05% Tween-20. Goat anti-rabbit antibody conjugated to a horseradish peroxidase decorated dextran-polymer backbone (Envision, DAKO Corp.) was used as a secondary reagent. For manual analysis, slides were visualized with diaminobenzidine (DAB, DAKO Corp.), followed by ammonium hydroxide acidified hematoxylin. For automated analysis, the primary antibody was co-incubated with monoclonal anti-cytokeratin (1:200, AE1/AE3, DAKO Corp.). Alexa-488-conjugated goat anti-mouse (Molecular Probes, Eugene, OR) was co-incubated with the secondary antibody. DAPI was added to visualize nuclei. For automated analysis Her-2 was visualized with a fluorescent chromagen (Cy-5-tyramide, NEN Life Science Products, Boston, MA) which, like DAB, is activated by HRP and results in the deposition of numerous covalently associated Cy-5 dyes immediately adjacent to the HRP-conjugated secondary antibody. Cy-5 (red) was used because is emission peak is well outside the green-orange spectrum of tissue autofluorescence.

Automated image acquisition and analysis

Automated image acquisition and analysis using AQUA has been previously described [4]. In brief, monochromatic, high-resolution (1024x1024 pixel, 0.5 micron resolution) images were obtained of each histospot. Areas of tumor were distinguished from stromal elements by creating a mask from the cytokeratin signal. Coalescence of cytokeratin at the cell surface helped localize the cell membranes, and DAPI was used to identify nuclei. HER2 signal from the membrane area of tumor cells was scored on a scale of 0-255, and expressed as signal intensity divided by the membrane area.

Fluorescence In-situ Hybridization

FISH analysis was performed with the PathVysion™ HER-2 DNA Probe Kit (Vysis, Downers Grove, IL) using two directly labeled fluorescent DNA probes complementary to the HER-2/*neu* gene locus (LSI HER-2/*neu* SpectrumRed™) and to chromosome 17 peri-centromeric _ satellite DNA (CEP17 SpectrumGreen™). The tissue array was baked at 60°C overnight, deparaffinized in xylene, and rinsed in 100% ethanol. The array was then pretreated in sodium thiocyanate (Vysis) at 80°C for 10 minutes, digested with pepsin (4 mg/ml in 0.01 M HCl; Vysis) for 15 minutes, and rinsed in graded ethanols (70%, 85%, and 99%). Twenty _l of probe/hybridization mixture including blocking DNA were applied to cover the entire array, covered with a 22 x 40 mm cover slip, and sealed with rubber cement. The probes and tissues were codenatured and hybridized using the Vysis Hybrite hybridization system. The Hybrite unit was programmed to allow 5 minutes of denaturation at 73°C, followed by overnight hybridization at 37°C. The array was washed in 2X SSC/.3% NP-40 at room temperature to remove the cover slip, and for 2 additional minutes at 73°C. After air drying the array was counter stained with 4',6-diamindino-2-phenylindole (DAPI) and cover-slipped.

HER-2/*neu* gene amplification was quantified by comparing the ratio of LSI HER-2/*neu* to CEP 17 probe signals in accordance with the PathVysion HER-2 DNA Probe kit criteria. All sections were examined directly using an Olympus AX70 epifluorescence microscope equipped with narrow band pass filters. Each histospot on the array was initially scanned at low power to assess heterogeneity and identify appropriate areas of tumor tissue with clearly defined, non-overlapping nuclei. The 60X objective was then used to score signals in 60 non-overlapping tumor cell nuclei to determine the average number of HER-2/*neu* and chromosome 17 copies/cell for each tissue specimen. The ratio of these averages was used to determine the presence of HER-2/*neu* gene amplification. Specimens with a HER-2/*neu*:chromosome 17 ratio greater than two were scored as positive for HER-2/*neu* gene amplification..

Data Analysis

Manual scoring of HER2 expression was assessed by a pathologist (R.L.C.) using a nominal four-point scale (0 to 3+). Histospot containing <10% tumor, as assessed either subjectively (manual) or by mask area (automated), were excluded from further analysis. Our previous studies have demonstrated that the HER2 score from a single histospot is representative of the score from an entire tissue section >90% of the time[11]. Subsequent studies revealed that analysis of a single histospot could provide significant statistical power to judge outcomes (data not shown), so that reported scores represent the result of only one histospot. Correlations with other prognostic markers were determined using Chi-square analysis. Overall survival analysis was assessed using Kaplan-Meier analysis with the Mantel-Cox logrank score for determining statistical significance. Relative risk was assessed using the univariate and multivariate Cox-proportional hazards model. Analyses were performed using Statview 5.0.1 (SAS Institute, Cary NC). Patients were deemed "uncensored" if they died of breast cancer within 30 years of their initial date of diagnosis.

Results and Discussion

Validation of microarray cohort

To validate our tissue microarray cohort of 300 node-positive breast cancers, we assessed several traditional histopathologic markers of malignancy. Using univariate analysis of long-term disease related survival, we found that large tumor size, high nuclear grade, low estrogen receptor expression, and high number of involved lymph nodes were all significant predictors of poor outcome (Table I). Next, we assessed the prognostic power of HER2 immunohistochemistry, using standard brown staining, visual examination by a pathologist, and scoring on a four-point scale (0-3+). Manual analysis showed a typical pattern of HER2 expression with 15% of tumors overexpressing the antigen (2+ and 3+, Table II). As expected, high level (3+) tumors showed a significantly worse outcome with a relative risk of 2.25 ($p = 0.0007$, Table I). Analysis of HER2 gene amplification using FISH was not predictive in our study; however, this is most likely due to the relatively small number of cases that, for technical reasons, were scorable (Table 1). However, results from FISH analysis and manual immunohistochemistry were highly correlated, similar to published findings ($p < 0.0001$, Table 2) [12,13].

HER2 expression on normal epithelium

We then assessed the level of HER2 expression using automated analysis on a microarray of normal breast epithelium. This epithelium was derived from normal ducts and/or lobules isolated from uninvolved breast tissue taken from 84 breast cancer patients. Consistent with prior studies using biochemical assays, our results demonstrated a low but detectable level of HER2 in normal epithelium, which was tightly grouped into a single peak with a mean of 5 and a standard deviation of 1.5 (AQUA score, Fig. 1, panel A) [14].

Automated analysis of HER2 expression in breast cancer

In contrast to the tightly grouped peak in normal epithelium, HER2 expression in breast tumors was broadly distributed (Fig. 1, Panel C). Tumoral levels of HER2 exhibited a mode similar to that of normal epithelium, but with significant skew toward higher-level expression. Examination of the histogram suggested that there were three naturally-occurring populations based on HER2 expression: normal, intermediate, and high (Fig. 1, Panel C). A discernible break in the histogram at AQUA score 25 divided HER2-high from the remaining tumors. The remaining tumors could then be subdivided into HER2-low and HER2-intermediate groups depending on whether their expression levels were greater than the mean HER2 expression of normal epithelium plus one standard deviation (AQUA score < 6.5, Fig 1, panels A & C). Using these divisions, 17.5% of the tumors were designated HER2 normal, 71.3% and HER2 intermediate, and 11.2% as HER2 high.

Comparison of manual and automated techniques

We then compared HER2 expression as gauged using automated and manual techniques (Fig 1, panels C and B, respectively). Regression analysis demonstrated good correlation between the two methods ($R = 0.704$). However, there was a significant

degree of overlap in the automated scores of cases from adjacent manually determined groups (Fig 1, panel D). Whereas there was a clear division between the histograms of tumors scoring 0/1+ and 2+/3+, the distinction between tumors scoring 0 and 1+ was indistinct. This result evidences the difficulty in manually translating a continuous variable into a nominal four-point scale, and distinguishing between marker levels on the low end of the scale (e.g. between normal and intermediate expression).

Examination of manual and automated techniques revealed that both were equally able define a population of tumors expressing high-level HER2 with poor outcome (relative risks = 2.25 and 2.18, p-value = 0.0007 and 0.0013, respectively, Table I). Furthermore, both showed a high degree of correlation with HER2 gene amplification as assessed by FISH ($p < 0.0001$ for both, Table II). However, unlike manual analysis, automated analysis revealed that HER2-normal expressing tumors also showed a significantly worse outcome (relative risk = 1.71, p-value = 0.0091, Table I). Given the amount of overlap in the 0 and 1+ categories from manual scoring (Fig. 1, Panel D), it is not surprising that manual assessment of stained slides have not previously identified the HER2 normal population.

Defining the sub-population of HER2-normal tumors

To determine if HER2 expression correlated with known prognostic markers in our cohort, we assessed possible associations between HER2 and hormone receptor status, tumor size, and nuclear grade. As expected, high-level HER2 expression is correlated with high nuclear grade and inversely correlated with estrogen receptor status (Table 3). The HER2-normal population showed no significant correlation with nodal involvement, tumor size, or estrogen receptor, but did associate with high nuclear grade ($p = 0.0494$, Table III). Few of the HER2-normal tumors exhibited gene amplification (Table II), ruling out the possibility that the HER2 gene in normal level tumors is amplified but that the HER2 protein is not detected, as might occur if antibody-reactive epitopes were mutated.

Multivariate analysis of HER2 normal and high populations

Finally, we determined whether either the HER2 normal or HER2 high levels were independent predictors of long-term disease-related survival. Combined multivariate analysis of HER2 with the traditional histopathologic markers, nodal involvement, tumor size, nuclear grade, and estrogen receptor, demonstrated that both normal and high level HER2 expression were independently predictive of patient outcome (Table IV).

Summary

Our data suggest that HER2 divides cases of node-positive breast carcinoma into three categories, normal, intermediate, and high expressors. Tumors expressing either normal or high HER2 levels do poorly in long-term follow-up. Of particular note are several prior studies that have looked at HER2 expression levels using "gold standard" biochemical techniques (Western blots and ELISAs)[5-8,14]. Several of these studies have shown that low-level HER2 expression correlated with tumor aggression[5-7]; although one did not find such an association [8]. Because such techniques require fresh tissue for

analysis, they were unable to assess long-term follow up on a large cohort of patients is difficult. Our studies have benefited from the use of tissue microarrays using existing paraffin-embedded tissues, which allowed the simultaneous examination of HER2 levels on a large cohort of patients with known long-term disease-related survival. Our data show that normal-level HER2 expression is an independent prognostic indicator of poor outcome, and demonstrate that unlike manual immunohistochemical analysis, automated analysis can identify a patient population that is otherwise only detectable using established biochemical assays.

HER2 overexpression can induce an aggressive phenotype via the activation of downstream regulators (e.g. Phosphoinositol-3-kinase, *Erk*/MAP kinase, and *Ras*) [15-18]. How normal levels of HER2 could be associated with a similar aggressive phenotype is currently unknown. We speculate that these tumors might over-express another growth factor receptor that promotes tumor aggression via a ligand-depdendent or –independent mechanism. It is possible that expression of such alternate growth factor receptors in some tumors results in the downregulation of HER2 expression via a feedback mechanism, resulting in aggressive tumors bearing a HER2-normal phenotype. Another possible explanation for the poor prognosis of HER2-normal tumors is that high-level activation of HER2 might result in the internalization and degradation of the receptor, resulting in apparent low-level HER2 expression. Finally, HER2-normal breast cancers may represent a population of aggressive poorly-differentiated neoplasms that have developed HER2 and growth factor independent mechanisms for their growth. The association between normal HER2 expression levels and high nuclear grade support this idea.

From a clinical perspective, response to Herceptin has largely been seen in HER2 high expressors or HER2 amplified cases. This may be due to the fact that 2+ or 3+ levels of expression were required for entry into most clinical trials[19-21]. The response of 0 or 1+ tumors to paclitaxel with and without Herceptin is currently being studied in a large randomized trial (CALGB 9840) [13]. Paradoxically, given the poor survival for patients with HER2-normal tumors, such patients may actually benefit more from Herceptin therapy than patients with HER2-intermediate tumors. Since prior studies have assessed HER2 levels using manual techniques, any treatment benefit for HER2-normal cases would presumably have been hidden. Our studies suggest that accurate distinction between HER2-normal and HER2-intermediate tumors requires automated analysis.

Table 1: Univariate analysis of 30-year disease-related survival

Marker	p-value	Relative Risk	95% Confidence Interval
HER2 Manual Score	0.0071		
0		1.00	
1+	0.9383	1.02	0.68 – 1.52
2+	0.9763	1.01	0.49 – 2.08
3+	0.0007	2.25	1.41 – 3.58
HER2 AQUA Score	0.0009		
Intermediate		1.00	
High	0.0013	2.18	1.35 – 3.51
Normal	0.0091	1.71	1.14 – 2.56
HER2 Amplification (FISH)	0.8121	1.07	0.60 – 1.90
Nodal Involvement	0.0279		
1 to 3		1.00	
4 to 9	0.6708	1.08	0.75 – 1.55
10 or more	0.0086	1.62	1.13 – 2.33
Tumor Size	0.0007		
< 2 cm		1.00	
2 – 5 cm	0.1255	1.33	0.92 – 1.93
> 5 cm	0.0001	2.09	1.43 – 3.07
Nuclear Grade			
High	0.0040	1.55	1.15 – 2.08
Estrogen Receptor			
Negative	0.0262	1.41	1.041 – 1.906

Table 2: HER2 amplification

		Protein Level Expression		FISH Amplification*	
Manual Score	0	191 / 300	63.7 %	4 / 100	4.0 %†
	1+	65 / 300	21.7 %	7 / 16	43.8 %
	2+	15 / 300	5.0 %	5 / 7	71.4 %
	3+	29 / 300	9.7 %	9 / 12	75.0 %
AQUA Score	Normal	47 / 269	17.5 %	2 / 21	9.5 %†
	Intermediate	192 / 269	71.3 %	13 / 95	13.7 %
	High	30 / 269	11.2 %	7 / 9	77.8 %

*P-values for AQUA scores vs. FISH amplification and manual scores vs. FISH amplification are both < 0.0001

†The difference between 9.5% and 4.0% is not statistically significant (p = 0.570)

Table 3: Distribution of prognostic markers by Her-2 level – Chi-Square Analysis

Marker	All Cases	Normal (%)	Intermed. (%)	High (%)	Chi-square p-value	
					Normal vs. Intermediate	High vs. Intermediate
Nodes Positive	n = 268				0.8171	0.1891
1 to 3	54%	60%	54%	43%		
4 to 9	26	23	27	33		
10 or more	20	17	19	23		
Tumor Size	n = 238				0.3033	0.8911
< 2 cm	43	35	44	48		
2 – 5 cm	34	44	32	32		
> 5 cm	23	21	24	20		
Nuclear Grade	n = 269				0.0494	0.0011
High	36	45	30	60		
Estrogen Receptor	n = 263				0.5325	<0.0001
Negative	40	39	34	77		

Table 4: Multivariate analysis of 30-year disease-related survival Table 1: Univariate analysis of 30-year disease-related survival

Marker	p-value	Relative Risk	95% Confidence Interval
HER2	0.0097		
Intermediate		1.00	
High	0.0136	1.96	1.15 – 3.36
Normal	0.0191	1.68	1.09 – 2.59
Nodal Involvement	0.1058		
1 to 3		1.00	
4 to 9	0.5915	1.12	0.73 – 1.72
10 or more	0.0353	1.61	1.03 – 2.53
Tumor Size	<0.0001		
< 2 cm		1.00	
2 – 5 cm	0.2220	1.31	0.85 – 2.01
> 5 cm	<0.0001	2.59	1.67 – 4.02
Nuclear Grade	0.2158		
High		1.26	0.87 – 1.82
Estrogen Receptor	0.0032		
Negative		1.75	1.21 – 2.54

Camp et al. Figure 1

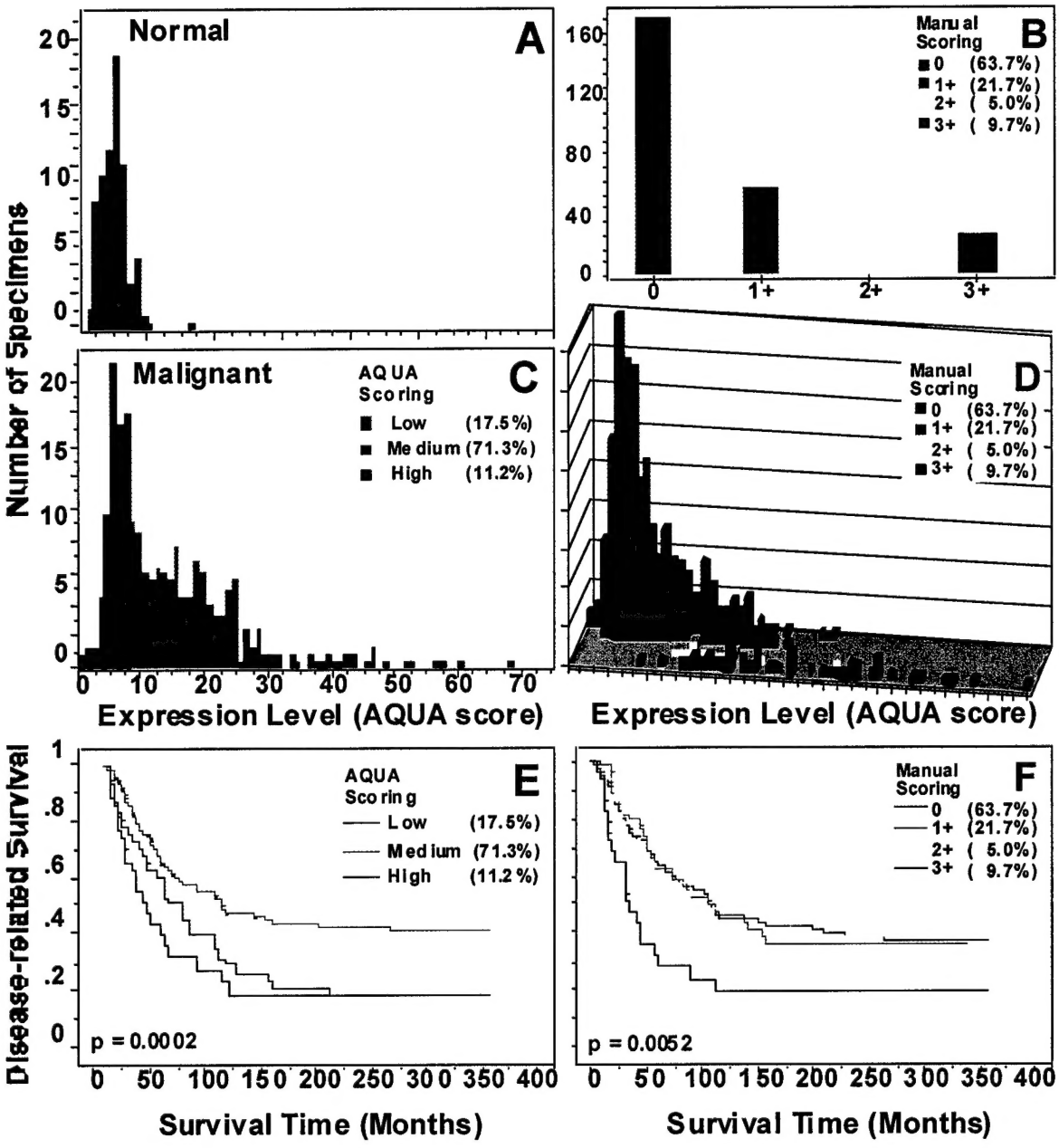

Figure 1: Automated analysis of HER2 divides tumors into three categories based on their level of expression. Panel A: Analysis of 84 samples of normal epithelium demonstrates a low but detectable level of HER2 expression. Panel B: Manual (visual) analysis of HER2 staining on a cohort of 300 node-positive carcinomas using a nominal four-point scale shows that 15% of the tumors over-express HER2 (2+/3+). Panel C: A histogram of HER2 expression is derived from automated (AQUA) analysis of HER2 expression on the same cohort. Cases are divided by expression level as follows: high (AQUA score > 25), normal (AQUA score < mean expression of normal epithelium + one standard deviation), and intermediate (between normal and high). Panel D: AQUA scores of tumors according to their manual score (0 – 3+) shows significant overlap, particularly between 0 and 1+ tumors. Panel E: Kaplan-Meier analysis of automated HER2 scores shows that both normal and high-level expressers do poorly relative to intermediate-level tumors. Panel F: Kaplan-Meier analysis of manual HER2 scores distinguishes a survival difference only with the high (3+) expressers.

Acknowledgements:

We thank Michael DiGiovanna and D. Craig Allred for critical review of this manuscript. We also thank Thomas D'Aquila, Mary Helie, Lori Charette, and Diana Fischer for their help in this effort. This work was supported by grants from the William and Catherine Weldon Donaghue Foundation for Medical Research and grants from the NIH including: K0-8 ES11571, NIEHS (RLC), RO-1 GM57604 NCI (DLR) and US Army DAMD grant 01-000436.

References:

1. Paik S, Bryant J, Tan-Chiu E, et al. Real-world performance of HER2 testing--ational Surgical Adjuvant Breast and Bowel Project experience. *J Natl Cancer Inst* 2002;**94**(11):852-4.
2. Roche PC, Suman VJ, Jenkins RB, et al. Concordance between local and central laboratory HER2 testing in the breast intergroup trial N9831. *J Natl Cancer Inst* 2002;**94**(11):855-7.
3. Zujewski JA. "Build quality in"--HER2 testing in the real world. *J Natl Cancer Inst* 2002;**94**(11):788-9.
4. Camp R, Chung G, Rimm D. Automated subcellular localization and quantitation of marker expression in tissue microarrays. *Nature Medicine* Submitted.
5. Dittadi R, Brazzale A, Pappagallo G, et al. ErbB2 assay in breast cancer: possibly improved clinical information using a quantitative method. *Anticancer Res* 1997;**17**(2B):1245-7.
6. Koscielny S, Terrier P, Spielmann M, Delarue JC. Prognostic importance of low c-erbB2 expression in breast tumors. *J Natl Cancer Inst* 1998;**90**(9):712.
7. Koscielny S, Terrier P, Daver A, et al. Quantitative determination of c-erbB-2 in human breast tumours: potential prognostic significance of low values. *Eur J Cancer* 1998;**34**(4):476-81.
8. Ferrero-Pous M, Hacene K, Tubiana-Hulin M, Spyratos F. Re: Prognostic importance of low c-erbB2 expression in breast tumors. *J Natl Cancer Inst* 1999;**91**(18):1584-5.
9. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens [see comments]. *Nat Med* 1998;**4**(7):844-7.
10. Katoh AK, Stemmler N, Specht S, D'Amico F. Immunoperoxidase staining for estrogen and progesterone receptors in archival formalin fixed, paraffin embedded breast carcinomas after microwave antigen retrieval. *Biotech Histochem* 1997;**72**(6):291-8.
11. Camp RL, Charette LA, Rimm DL. Validation of tissue microarray technology in breast carcinoma. *Lab Invest* 2000;**80**(12):1943-9.
12. Birner P, Oberhuber G, Stani J, et al. Evaluation of the United States Food and Drug Administration-approved scoring and test system of HER-2 protein expression in breast cancer. *Clin Cancer Res* 2001;**7**(6):1669-75.
13. Seidman AD, Fornier MN, Esteva FJ, et al. Weekly trastuzumab and paclitaxel therapy for metastatic breast cancer with analysis of efficacy by HER2 immunophenotype and gene amplification. *J Clin Oncol* 2001;**19**(10):2587-95.
14. Dittadi R, Donisi PM, Brazzale A, Marconato R, Spina M, Gion M. Immunoenzymatic assay of erbB2 protein in cancer and non-malignant breast tissue. Relationships with clinical and biochemical parameters. *Anticancer Res* 1992;**12**(6B):2005-10.
15. Dankort DL, Wang Z, Blackmore V, Moran MF, Muller WJ. Distinct tyrosine autophosphorylation sites negatively and positively modulate neu-mediated transformation. *Mol Cell Biol* 1997;**17**(9):5410-25.

16. Ben-Levy R, Paterson HF, Marshall CJ, Yarden Y. A single autophosphorylation site confers oncogenicity to the Neu/ErbB-2 receptor and enables coupling to the MAP kinase pathway. *Embo J* 1994;**13**(14):3302-11.
17. Amundadottir LT, Leder P. Signal transduction pathways activated and required for mammary carcinogenesis in response to specific oncogenes. *Oncogene* 1998;**16**(6):737-46.
18. Janes PW, Daly RJ, deFazio A, Sutherland RL. Activation of the Ras signalling pathway in human breast cancer cells overexpressing erbB-2. *Oncogene* 1994;**9**(12):3601-8.
19. Vogel CL, Cobleigh MA, Tripathy D, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J Clin Oncol* 2002;**20**(3):719-26.
20. Cobleigh MA, Vogel CL, Tripathy D, et al. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J Clin Oncol* 1999;**17**(9):2639-48.
21. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 2001;**344**(11):783-92.